EINFACHE DNA-SEQUENZEN

IN

EUKARYONTEN

Dissertation
zur Erlangung des Grades eines Doktors
der Naturwissenschaften

der Fakultät für Biologie
der Eberhard-Karls-Universität Tübingen

vorgelegt von

Diethard Tautz
aus Glonn/Obb.

1983

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547612_0003

EINFACHE DNA–SEQUENZEN

IN

EUKARYONTEN

Dissertation
zur Erlangung des Grades eines Doktors
der Naturwissenschaften

der Fakultät für Biologie
der Eberhard-Karls-Universität Tübingen

vorgelegt von

Diethard Tautz
aus Glonn/Obb.

1983

Tag der mündlichen Prüfung

14. Juli 1983

Dekan

Prof. Dr. K.P. Hadeler

1. Berichterstatter

Privatdozent Dr. M. Renz

2. Berichterstatter

Prof. Dr. W. Seyffert

UNIV.
LEIDEN
BIBL.

Diese Arbeit wurde in der Zeit vom Juni 1981 bis zum März 1982 am Max-Planck-Institut für Biologie in Tübingen, Abteilung Beermann und vom April 1982 bis Mai 1983 am Europäischen molekularbiologischen Labor (EMBL) in Heidelberg im Labor von Manfred Renz und Jan-Eric Edström angefertigt. Ich danke einer Reihe von Mitarbeitern beider Institute für ihre Hilfestellungen. Besonderer Dank gilt Jan-Eric Edström für die Hilfe bei der Mikroklonierung. Christina Kurz danke ich für die Hilfe beim Sequenzieren.

ABKÜRZUNGEN

DNA	Desoxyribonukleinsäure
RNA	Ribonukleinsäure
kb	Kilobasenpaare
nt	Nukleotide
cpm	counts per minute (Zählrate)
RT	Raumtemperatur (zwischen 18 - 22°C)
DH	Denhardt´s Lösung nach (Denhardt, 1966)
EDTA	Ethylendiamintetraessigsäure, mit NaOH auf pH 8,0 titriert
EtBr	Ethidiumbromid
DEPC	Diethylpyrocarbonat
SDS	Natriumdodecylsulfat
Tris-HCl	Tris(hyroxymethyl)-aminomethan mit HCl auf den jeweils angegebenen pH-Wert titriert
MOPS	3-Morpholinopropansulfonsäure

Einfache DNA-Sequenzen wurden wie folgt abgekürzt:
Bei Homopolymeren wurden jeweils zwei Nukleotide von jedem Strang geschrieben, also z.B. AA/TT statt poly dA/poly dT. Bei Kopolymeren mit einer Repetetionseinheit von zwei oder mehr Nukleotiden, wurde jeweils die Repetetionseinheit von beiden Strängen geschrieben also z.B. GT/CA statt poly (dG-dT)/poly(dC-dA).

Inhaltsverzeichnis

		Seite
EINLEITUNG		7
MATERIAL UND METHODEN		
1.	Material	11
2.	Isolierung von DNA und RNA über CsCl-Gradienten	12
3.	Mikroisolierung von RNA aus Drosophila Geweben	14
4.	Fraktionierung von Kc-Zellen	15
5.	DNA und RNA Gele / DNA Markerfragmente	16
6.	Blotting von DNA und RNA / Färbung von RNA-Filtern	17
7.	Hybridisierung / Kalkulation der Bedingungen	18
8.	In situ Hybridisierung	22
9.	Enzymreaktionen	22
10.	Klonierung	24
11.	Sequenzierung	25
ERGEBNISSE		
I.	Simple Sequenzen in der DNA	26
I.1.	Nachweis durch dot-blot Hybridisierungen	26
I.2.	Quantifizierung	29
I.3.	Genomblothybridisierungen	31
I.4.	In situ Hybridisierungen	33
II.	Simple Sequenzen in der RNA	35
II.1.	Hybridisierung an Gesamt-RNA	35
II.2.	Hybridisierung an fraktionierte Kc-Zell RNA	36

Seite

II.3. Hybridisierung an stadien- und gewebe-spezifische Drosophila RNA 39

III. Klonierung und Sequenzierung 42

III.1. Isolierung und Charakterisierung von Klonen aus zwei Mikroklonbänken 44

III.2. Klone, die für die Sequenzierung ausgewählt wurden 49

III.3 Strategie für die Restriktionskartierung 50

III.4 Restriktionskarten und Sequenzen 54

III.4.1. Klon pDv-1 54

III.4.2. Klon pDv-142 58

III.4.3. Klon pDv-161 60

III.4.4. Klon pDv-191 65

III.4.5. Referenzklone pDv-182 und pMm-2 68

III.5. Zusammenfassung der Sequenzierungsergebnisse 70

III.6. Genomblothybridisierungen an D.virilis DNA 72

III.6.1. Hybridisierung mit Klon pDv-192 72

III.6.2. Hybridisierung mit Fragmenten aus Klon pDv-161 73

IV. Diskussion der Experimente 76

IV.1. Hybridisierungsspezifität 76

IV.2. Klonierungsbedingungen 80

DISKUSSION 81

ZUSAMMENFASSUNG 90

LITERATUR 91

EINLEITUNG

Eukaryontische Genome enthalten neben Sequenzen, die nur ein, oder wenige Male vorkommen (z.B. die meisten Gene) unterschiedliche Klassen von repetetiven Sequenzen (Britten & Kohne, 1968). Mit Hilfe von Renaturierungsstudien lassen sich diese grob in zwei Klassen einteilen. Die Klasse der hochrepetetiven Sequenzen mit einem Repetetionsgrad von über 10^6, sowie die Klasse der mittelrepetetiven mit einem Repetetionsgrad von 10^2 bis 10^5 (Britten & Kohne, 1968). Bei den hochrepetetiven Sequenzen (auch Satelliten-DNA genannt, da sie im Dichtegradienten oft an einer anderen Stelle bandet als der Rest der DNA) handelt es sich in der Regel um kurze tandemartig viele tausendmal wiederholte Sequenzen. Die Länge der sich wiederholenden Einheiten variiert dabei zwischen zwei bis einige hundert Nukleotide (aber auch noch längere Einheiten wurden gefunden), die dadurch oft sehr lange (> 20 kb) einfache Sequenzen bilden (Brutlag, 1980). Über ihre Entstehung und Funktion ist wenig bekannt, aber es gibt eine Reihe von Hypothesen dazu (zusammenfassend dargestellt in Brutlag, 1980).

Die mittelrepetetiven Sequenzen bilden eine sehr heterogene Klasse. Man kann sie noch einmal unterteilen in solche Elemente, die relativ kurz sind (< 500 bp, 'SINEs'), sowie in solche, die relativ lang sind (> 5 kb, 'LINEs') (Singer, 1982). Die Klasse der SINEs, zu der zum Beispiel auch die Alu-Elemente beim Mensch (Schmid & Jelinek, 1982) gehören, ist relativ gut charakterisiert. Diese Elemente sind verstreut im Genom angeordnet, besitzen oft einen Promotor

für RNA-Polymerase III (Jagadeeswaran et al., 1981) und werden in der Regel auch in der RNA gefunden. In der Kern-RNA treten sie oft als Teile von längeren Transkripten, die selbst nicht repetetiv sind, auf (Davidson & Posakony, 1982). Während der RNA-Prozessierung gehen sie größtenteils verloren, aber in einigen Fällen konnten sie auch in cytoplasmatischer mRNA nachgewiesen werden (Georgiev et al., 1981; Calabretta et al., 1981). An den Grenzen enthalten diese Elemente oft einige Nukleotide, die sich vor und nach dem Element direkt wiederholen (Singer, 1982). Solche direkten repeats findet man auch bei transponierbaren Elementen in Prokaryonten (Calos & Miller, 1980). Es wird daher spekuliert, ob diese Elemente auch in Eukaryonten durch eine Art Transposition entstehen, die aber, im Gegensatz zu Prokaryonten, über die Reintegrierung einer cDNA von einem Transkript dieses Elements laufen würde (Jagadeeswaran et al., 1981; Sharp, 1983). In mindestens einem Fall ließ sich solch ein Transpositionsereignis auch zeigen (Kominami et al., 1983), wobei gefunden wurde, daß die direkten repeats an den Enden, ähnlich wie bei Prokaryonten, durch eine Verdoppelung der genomischen DNA an dieser Stelle entstanden sind.

Diese Arbeit befaßt sich mit einer anderen Klasse von mittelrepetetiven Sequenzen, nämlich simplen Sequenzen. Es gab schon früher Hinweise, daß simple Sequenzen, die aus Homopolymeren eines Nukleotides oder aus Kopolymeren von zwei alternierenden Nukleotiden bestehen, in eukaryontischen Genomen vorkommen. Zum Beispiel wurde die Sequenz AA/TT (siehe 'Abkürzungen') in allen daraufhin untersuchten Eukaryonten gefunden und auch eingehend untersucht (Shenkin & Burdon, 1972; Bishop et al., 1974; Mol et al., 1976; Flavell

et al., 1977). In der Maus bildet sie Homopolymere von 15 bis 25 bp Länge (Flavell et al., 1977). Man findet diese Sequenz aber auch als oligo-U oder oligo-A Stücke in der Kern-RNA (Molloy et al., 1972; Dubroff & Nemer, 1975). Spätere Untersuchungen konnten sie auch in, an Polysomen gebundener mRNA nachweisen (Katinakis & Burdon, 1981). In Dictyostelium gibt es am 3'-Ende der meisten mRNAs neben der typischen poly-A Sequenz noch eine oligo-A Sequenz, die im Gegensatz zur ersteren offenbar durch Transkription entstanden ist (Jacobson et al., 1974) und die mit der Transkriptionstermination in Zusammenhang gebracht wird.

Es gibt aber noch mehr Hinweise auf simple Sequenzen. Eine Reihe von Arbeiten befaßt sich zum Beispiel mit Polypyrimidinsequenzen, die durch Depurinisierung mit Säure aus der DNA von Eukaryonten isoliert werden können (Straus & Birnboim, 1974; Birnboim & Straus, 1975). Diese sind 20 bis 200 nt lang und offenbar über das ganze Genom verteilt (Birnboim et al., 1979). Sequenzen, die aus unterschiedlichen Spezies isoliert werden, zeigen starke Sequenzhomologien in Kreuzhybridisierungsstudien (Straus & Birnboim, 1976). Außerdem sind sie auch in der RNA vertreten (Straus & Birnboim, 1976). Wahrscheinlich handelt es sich bei diesen Befunden ebenfalls um simple Sequenzen, zum Beispiel die Sequenz GA/CT oder aber auch solche höherer Ordnung, die aus mehr als zwei sich wiederholenden Nukleotiden aufgebaut sind, zum Beispiel GAA/CTT oder GGAA/CCTT. Letztere wurde auch tatsächlich in einer klonierten Polypyrimidin DNA gefunden (Deugau et al., 1983). Die einfache Sequenz GA/CT wurde in, bzw. in der Nähe von Genen gefunden, zum Beispiel in einem nicht transkribierten

Teil des Seeigel Histongenclusters (Schaffner et al., 1978; Sures et al., 1978), oder in einem Intron des Histokompatibilitätsgens H-2k^{d} der Maus (Kvist et al., 1983). Durch Sequenzierung wurde auch die einfache Sequenz GT/CA gefunden, zum Beispiel in der Nähe des 3´-Endes des κ V-Region Immunglobulingens der Maus (Nishioka & Leder, 1980) oder in einem Intron des menschlichen γ- Globin Gens (Slightom et al., 1980), sowie auch noch in einer Reihe anderer Fälle (zusammengefaßt in Nordheim & Rich, 1983). Kürzlich konnte gezeigt werden, daß diese einfache Sequenz mittelrepetetiv und im Genom aller daraufhin untersuchten Eukaryonten verteilt ist (Hamada et al., 1982b). Ähnliches gilt auch für die simple Sequenz GC/CG. Schon früher wurde auch die einfache Sequenz GG/CC als mittelrepetetive Komponente von einigen Mammaliern beschrieben (Shenkin & Burdon, 1974).

In dieser Arbeit wurde nun eine systematische Suche nach einfachen Sequenzen in Eukaryonten durchgeführt. Es konnte gezeigt werden, daß alle sechs möglichen Kombinationen aus nur einem, oder zwei alternierenden Nukleotiden, nämlich AA/TT,GG/CC,AT/TA,GC/CG,GT/CA und GA/CT als mittelrepetetive Komponenten in eukaryontischen Genomen vorkommen. Im Laufe der Arbeit wurden auch simple Sequenzen gefunden, die aus drei tandemartig wiederholten Nukleotiden bestehen, nämlich CAG/GTC und GTT/CAA. Noch weitere simple Sequenzen höherer Ordnung wurden auch von anderen Autoren gefunden (Höchtl & Zachau, 1983; Epplen et al., 1982). Weiterhin konnte in dieser Arbeit gezeigt werden, daß simple Sequenzen auch in RNA transkribiert werden und zum Teil sogar in poly-A^{+} und cytoplasmatischer RNA vorkommen.

MATERIAL UND METHODEN

1. Material

Für die Isolierung von DNA und RNA aus den verschiedenen Eukaryonten wurde folgendes Ausgangsmaterial benutzt:

Mensch: Eine primäre Zellkultur aus menschlichen Fibroblasten (erhalten von Dan Röhme, EMBL Heidelberg)

Xenopus: Leber von Xenopus laevis (erhalten von Peter Hausen, MPI für Virusforschung, Tübingen)

Drosophila: Zellkultur einer embryonalen Zellinie von D.melanogaster (Kc-Zellen). Die Zellen wurden in D 20 Medium nach Echalier und Ohanessian (1970) in Suspensionskultur gezogen. Weiterhin wurden noch Fliegen und Larven der Art D.virilis (erhalten von W. Beermann, MPI für Biologie Tübingen) verwendet, die auf dem Standard-Maisbrei gezogen wurden. RNA aus verschiedenen Stadien von D.melanogaster wurde erhalten von Herrmann Steller, EMBL Heidelberg.

Seeigel: DNA wurde gewonnen aus Seeigel Spermien (erhalten von Hans Walter Zentgraf, DKFZ Heidelberg), RNA aus unbefruchteten Seeigeleiern (erhalten von Christian Petzelt, DKFZ Heidelberg).

Weizen: Grüne Blätter einer Winterweizensorte (Triticum aestivum) von einem Feld bei Heidelberg.

Hefe: Hefe-Zellen aus der frühen Wachstumsphase (erhalten von Giovanni Cesareni, EMBL Heidelberg).

Synthetische DNA mit simpler Sequenz, poly-rA sowie nicht markierte Nukleotide stammten von der Firma Boehringer. Die oligo-Nukleotide $p(dAdG)_5$ und $p(dCdT)_5$ stammten von der Firma P.L. Biochemicals (Wisconsin). Radioaktive Nukleotide mit spezifischen Aktivitäten von 400 bis 4000 Ci/mmol stammten von der Firma Amersham.

Enzyme wurden bei Boehringer oder BRL gekauft. Ligase und Eco RI stammten von V. Pirotta, EMBL Heidelberg.

Alle nicht näher bezeichneten Chemikalien kamen von der Firma Merck.

Als Klonierungsvektoren wurden das Plasmid pBR 322 (Sutcliffe, 1978) und die λ-Phagen λ 641 (Murray et al., 1977) und λ EMBL 2 (abgeleitet von λ 1059 (Karn et al., 1980) von H. Lehrach, EMBL Heidelberg) verwendet. pBR 322 wurde in dem Bakterienstamm E.coli HB 101 gezogen. Die Phagen wuchsen in den E.coli Stämmen Q 358 (λ 641 und λ EMBL 2), Q 359 (zur Selektion von Inserttragenden Phagen des Stamms λ EMBL 2) und POP 101 (zur Selektion von Inserttragenden Phagen des Stamms λ 641). Die Beschreibung und Genetik dieser Stämme ist in Murray (1983) zusammengefaßt.

2. Isolierung von DNA und RNA über CsCl-Gradienten

DNA und RNA aus den oben genannten eukaryontischen Geweben wurde jeweils aus dem gleichen Ansatz auf folgende Weise isoliert. Das Gewebe wurde lysiert und wenn notwendig auch homogenisiert in 4M Guanidin-HCl, 1 M 2-Mercaptoethanol, 50

mM NaAcetat pH 6,0, 10 mM EDTA, 0,5% Sarkosyl (von Ciba Geigy). (Die Hefezellen wurden für 30 min vorlysiert in 1% SDS, 20 mM Tris-HCl pH 7,0 bei 65°C). In jeweils 3, 5 ml dieses Homogenats wurden 2,0g festes CsCl gelöst. Diese Lösung wurde dann in einem Nitrocelluloseröhrchen (0,5 x 2 inch) auf ein Kissen von 1 ml 5,7 M CsCl, 50 mM NaAcetat pH 6,0, 10 mM EDTA geschichtet.
Dieser Stufengradient wurde bei 36 000 rpm, 25°C im SW 50.1 Rotor (Beckmann) für 24h zentrifugiert. Dabei wird die RNA pelletiert, während die DNA im mittleren Drittel des Gradienten verbleibt (sie bildet dort noch keine Bande!). Die Gradienten wurden dann von oben per Hand fraktioniert. Die viskösen Fraktionen (enthalten die DNA) wurden gegen TE-Puffer (10 mM Tris-HCl pH 8,0, 1 mM EDTA) dialysiert, mit RNAase und α-Amylase extensiv verdaut, vorsichtig phenolisiert (2 x mit je 1 Vol: 80% Phenol in Wasser, 0, 1% Hydroxychinolin, mit Tris-OH neutralisiert), 2 x mit Chloroform und 2 x mit Ether nachextrahiert und nochmals gegen TE-Puffer dialysiert.

Die RNA aus dem Pellet der obigen Gradienten wurde in RNAase-freiem (siehe unten) TES-Puffer (10 mM Tris-HCl pH 7,0, 1 mM EDTA, 0,5% SDS) aufgenommen und 5 min auf 65°C erhitzt (vgl. Glišin et al., 1974). Danach wurde 1/10 Vol 3 M NaAcetat pH 5,0 und 2,5 Vol Ethanol zugegeben. Diese Mischung wurde für 10 min bei -70°C (Trockeneis) gefällt und 10 min abzentrifugiert (12 000g). Das Pellet wurde 2 x mit 80% Ethanol und einmal mit 100% Ethanol nachgewaschen im Vakuum getrocknet und dann in TES-Puffer oder in sterilem bidest. Wasser aufgenommen.

Die so gewonnene RNA ist in der Regel nicht degradiert und frei von DNA. Lediglich kleine RNA (< 5S) sedimentiert unter den obigen Bedingungen nicht vollständig und ist daher in der Gesamt-RNA Präparation unterrepräsentiert.

3. Mikroisolierung von RNA aus Drosophila Geweben

Das Verfahren zur Mikrosiolierung von RNA ist angelehnt an das oben beschriebene. Die isolierten Organe werden direkt in eine kleine Menge (≃ 30µl) des oben angeführten Lysepuffers gebracht. Sie lösen sich darin in ca. 5 min selbst auf, so daß keine Homogenisierung notwendig ist. Zu jeweils 34µl dieser Lösung werden dann 136µl einer 7,1 M CsCl Lösung gegeben. Das ganze wird dann in Nitrocelluloseöhrchen in der Airfuge (Beckmann) bei 25 psi über Nacht zentrifugiert.
Bei diesen Gradienten schwimmt am Ende die DNA und das Protein an der Oberfläche und kann teilweise mit der Pinzette entfernt werden. Das RNA Pellet befindet sich seitlich am Boden (Festwinkelrotor!). Von ca. 10 Drüsen aus D.virilis erhält man bereits ein sichtbares Pellet. Die überstehende Flüssigkeit wird vorsichtig mit einer Pipette entfernt. Die Wände des Röhrchens werden dann mit sterilem Zellstoff abgewischt. Das Pellet wird schließlich in TES-Puffer aufgenommen und gefällt wie oben beschrieben.

4. Fraktionierung von Kc-Zellen

Alle im folgenden verwendeten Puffer wurden zunächst RNAase-frei gemacht. Dazu wurden sie zunächst steril filtriert, dann mit Diethylpyrocarbonat (von Sigma) in einer Endkonzentration von 0,5% behandelt und schließlich autoklaviert.

Die Kc-Zellen wurden aus der Suspensionskultur abzentrifugiert (2000 rpm, HB4 Rotor von Sorvall) und einmal mit eiskaltem Puffer (100 mM NaCl, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 10 mM Tris-HCl pH 7,2) gewaschen. Alle weiteren Schritte wurden bei 0^oC und mit vorgekühlten Puffern und Gefäßen durchgeführt. Die Zellen wurden lysiert in 100 mM NaCl, 5 mM $CaCl_2$, 5 mM $MgCl_2$, 30 mM Tris-HCl pH 8,3, 0,5% DEPC, 0, 5% Nonidet P 40 (die beiden letzten Komponenten wurden erst kurz vor Gebrauch des Puffers zugegeben) und mit dem L-Stempel im Glashomogenisator vorsichtig homogenisiert.

Die Kerne wurden bei 3000 rpm (HB4-Rotor) abzentrifugiert, in Waschpuffer (100 mM NaCl, 1,5 mM EDTA, 10 mM Tris-HCl pH 8,0, 0, 5% DEPC) aufgenommen, mit dem L-Stempel homogenisiert und wieder abzentrifugiert. Dieser Waschschritt wurde nochmals wiederholt. Danach wurde die RNA wie in 2. beschrieben isoliert.

Der erste cytoplasmatische Überstand wurde nochmals bei 10 000 rpm (HB4-Rotor) zentrifugiert. Dieser Überstand wurde dann 1 : 1 mit Guanidin-HCl Puffer (s. 2.) gemischt und die RNA wurde wie in 2. beschrieben isoliert.

Die Fraktionierung in poly-A^+/poly-A^- RNA wurde im

wesentlichen nach Aviv und Leder (1972) durchgeführt. Die RNA wurde auf die oligo-dT Cellulose Säule in 0,5 M NaCl/TES aufgetragen, mit 0,1 M NaCl/TES wurde nachgewaschen und mit auf 50°C erwärmten TES-Puffer wurde die RNA eluiert.

5. DNA und RNA Gele

DNA wurde in der Regel auf Agaroseflachgelen (zwischen 0, 5 bis 1,2% Agarose) in Tris-Acetat Puffer (40 mM Tris-Acetat pH 8,1, 2 mM EDTA) nach Loening (1967) aufgetrennt, mit 10µg/ml EtBr in H_2O gefärbt und mit UV-Licht (366 oder 254 nm) sichtbar gemacht. Die DNA Proben wurden vor dem Auftragen mit 1/4 Vol 40% Glycerin, 20% Ficoll 400, 0,2% SDS, 0, 01% Bromphenolblau gemischt.

In der Regel wurden zwei verschiedene DNA Größenmarker auf die Gele mit aufgetragen. Marker 1: Phage λ (WT) mit Hind III geschnitten, 8 Fragmente mit folgenden Größen (in kb): 23,15, 9,42, 6,56, 4,38, 2,32, 2,02, 0,56 und 0,125. Marker 2: Plasmid pBR 322 geschnitten mit Pst I, Ava I, Eco RI und Bam HI, vier Fragmente mit folgenden Größen (in kb): 2,19, 1,05, 0,75 und 0,375.

RNA wurde auf 1,2% Agaroseflachgelen, die 2,2 M Formaldehyd enthielten (nach Lehrach et al., 1977) aufgetrennt. Der Gelpuffer und Elektrophoresepuffer war 20 mM MOPS, 5 mM NaAcetat, 1 mM EDTA mit NaOH auf pH 7,0 eingestellt.

Vor dem Auftragen wurde die RNA im Verhältnis 1 : 4 mit Probenpuffer (50% Formamid, 2, 2 M Formaldehyd in Elektrophoresepuffer) versetzt und 5 min bei 65°C

denaturiert. Zum Auftragen wurde danach noch etwas Farbmarker (0.01% Bromphenolblau) zugegeben.

6. Blotting von DNA und RNA

DNA aus Agarosegelen wurde im wesentlichen nach der von Southern (1975) beschriebenen Prozedur auf Nitrocellulosefilter geblottet. RNA aus Formaldehydgelen wurde in ähnlicher Weise geblottet. Diese wurden nach der Elektrophorese für 10 min in 20 x SSC (s.u.) geschüttelt und dann in 10 x SSC auf Nitrocellulose geblottet.
Nach dem Blotten wurden die Filter kurz in 1 x SSC (DNA) oder in 3 x SSC (RNA) gewaschen, auf Filterpapier abgetupft und im Vakuum für 2h bei 80°C gebacken.
Wenn ein RNA-Filter gefärbt werden sollte, wurde er zunächst für 15 min in 5% Essigsäure geschüttelt, dann für 15 min in 0, 04% Methylenblau in 0,5 M NaAcetat pH 6,0 gefärbt und in Wasser entfärbt.

Dot blots wurden mit Hilfe des Minifold-Apparates von Schleicher & Schuell hergestellt. Die Puffer- und Denaturierungsbedingungen entsprachen den von Kafatos et al. (1979) beschriebenen. Zur Quantifizierung der Hybridisierungssignale mit Hilfe von dot blots (vgl. I.2.) wurden jeweils drei 1 : 10 Verdünnungen der Proben (ausgehend von 1µg) sowie der Vergleichsplasmide (ausgehend von 0, 2µg) aufgetragen. Alle Verdünnungen wurden jeweils mit E.coli DNA auf 1µg DNA eingestellt (Endvolumen bei allen Proben: 0, 5 ml). Auf diese Weise können die Hybridisierungssignale auf

dem Röntgenfilm mit ausreichender Genauigkeit verglichen werden (Kafatos et al., 1979).

7. Hybridisierung

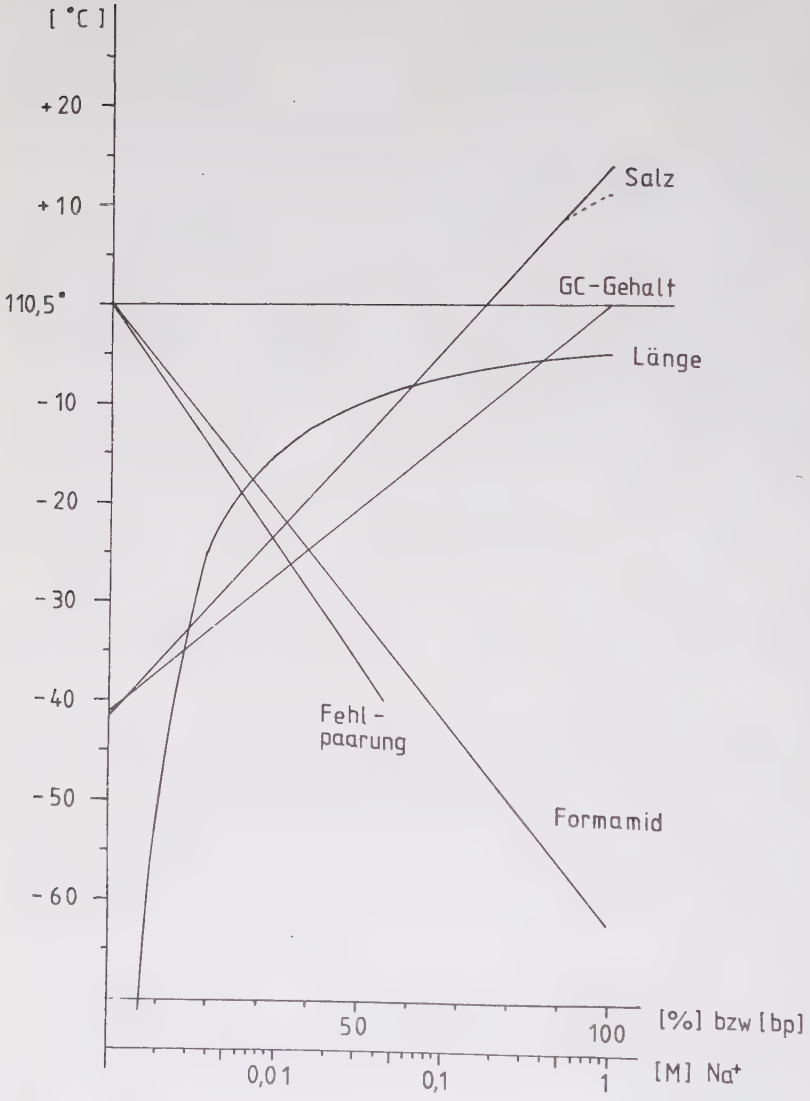
Die Vorhybridisierung erfolgte bei allen Filtern in 2 x Denhardt's Lösung (Denhardt, 1966) 5 x SSC (1 x SSC = 0,15 M NaCl, 0, 015 M Na-citrat), 0,5% SDS bei 65°C für 2 bis 5h. Die Hybridisierungslösungen enthielten neben unterschiedlichen Mengen SSC und Formamid (deionisiert, vgl. Casey & Davidson, 1977) 1 x DH, 0,5% SDS und in der Regel 20µg/ml E.coli DNA (denaturiert und geschert). Die radioaktiven Proben wurden in 95% Formamid aufgenommen, kurz auf 100°C erhitzt und heiß zugegeben. Damit wurde verhindert, daß die simplen Sequenzen zu schnell renaturierten. Hybridisiert wurde meistens über Nacht im Schüttelwasserbad in Volumina in denen sich die Filter frei bewegen konnten (zwischen 10-50ml).

Bei der Hybridisierung mit den simplen Polymeren mußten zwei Dinge berücksichtigt werden, nämlich erstens der unterschiedliche GC-Gehalt und zweitens, daß auch kurze, aber nicht zu kurze Stücke hybridisieren sollten. Beides hat einen Einfluß auf die Schmelztemperatur der Hybride. Diese ist aber auch noch von zwei weiteren Faktoren abhängig, nämlich der Salzkonzentration und der Formamidkonzentration. Man kann nun die Hybridisierungsbedingungen so kalkulieren, daß Hybride einer bestimmten Länge nicht mehr entstehen können, da ihre Schmelztemperatur unter der Hybridisierungs-

temperatur liegen würde. Hybride, die etwas länger sind, können dann zwar entstehen, aber nur mit einer geringeren Rate. Das Optimum für die Hybridisierungsrate liegt nämlich weit von der Schmelztemperatur entfernt (ca. 25°C darunter (Britten et al., 1974)). Auch die Salzkonzentration hat einen Einfluß auf die Hybridisierungsrate, wobei gilt, daß eine hohe Konzentration die Rate wesentlich steigert (Britten et al., 1974). Wenn man die Hybridisierungsbedingungen so kalkuliert, daß auch sehr kurze Sequenzen noch hybridisieren sollen, läuft man Gefahr, daß andererseits lange Sequenzen mit hohem Fehlpaarungsanteil auch hybridisieren, also unspezifische Signale entstehen. Dem kann man teilweise dadurch vorbeugen, daß man nach der Hybridisierung

Abb. 1: Abhängigkeit der DNA Schmelztemperatur von verschiedenen Parametern

Die Formeln für die einzelnen in der Graphik eingetragenen Parameter sind in Chow & Broker (1981) zusammengestellt. Für GC-Gehalt, Formamid und Fehlpaarung gilt die %-Skala, für die Länge die bp-Skala und für Salz die logarithmische Skala. Der Nullpunkt der Ordinate entspricht der Schmelztemperatur (unter Standardbedingungen, 0,18 M Na^+) von DNA, die nur aus GC-Nukleotiden besteht. Diese beträgt 110,5°C. Von diesem Wert müssen für andere DNAs, bzw. für andere Bedingungen, alle aus der Graphik ermittelten Werte (in °C) abgezogen (bzw. addiert) werden, um die zugehörige Schmelztemperatur zu erhalten. Die Salzabhängigkeit ist bei hohen Konzentrationen (>0,5 M) nicht mehr linear, es gilt dann eher die gestrichelte Linie.


[°C]
+20
+10
110,5°
-10
-20
-30
-40
-50
-60
Salz
GC-Gehalt
Länge
Fehl-
paarung
Formamid
50
100
[%] bzw [bp]
0,01
0,1
1
[M] Na+

unterschiedlich stringente Waschschritte durchführt, bei denen fehlgepaarte Bereiche (aber auch sehr kurze Sequenzen) wieder abgeschmolzen werden.
Der Einfluß der einzelnen Effekte auf die Schmelztemperatur kann mit Formeln kalkuliert werden und ist in Abb. 1 zusammenfassend dargestellt. Alle Effekte können einfach summiert werden, um die endgültigen Bedingungen zu erhalten.

Für die einzelnen Hybridisierungen wurden die folgenden Bedingungen gewählt:
Hybridisierung mit GT/CA und GA/CT (je 50% GC-Gehalt) sowie mit nicht simplen Sequenzen: 5 x SSC, 50% Formamid, 37°C.
Hybridisierung mit AA/TT und poly-rA (je 0% GC-Gehalt): 2 x SSC, 5% Formamid, 37°C.
Hybridisierung mit GG/CC (100% GC-Gehalt): 2 x SSC, 70% Formamid, 37°C.
Bei der Hybridisierung mit AT/TA und GC/CG unter den kalkulierten Bedingungen wurden keine vernünftigen Resultate erhalten. Da diese beiden Polymere selbstkomplementär sind, spielen hier wahrscheinlich noch weitere Effekte eine Rolle. GC/CG wurde daher in 5 x SSC, 50% Formamid bei 37°C hybridisiert, AT/TA in 2 x SSC, 5% Formamid bei 20°C (RT).

Nach der Hybridisierung wurde in der Regel zunächst für 2h mit der Hybridisierungslösung (ohne DH und ohne DNA) bei der Hybridisierungstemperatur nachgewaschen und dann 3 x in großem Volumen (ca. je 200 ml) 2 x SSC bei RT (= nicht stringenter Waschschritt). Als stringenter Waschschritt bei einer Reihe von Experimenten wurde dann noch 2 x in 0,1 x SSC bei 50°C gewaschen. (Der stringente Waschschritt für AA/TT

bzw. poly-rA war 0, 1 x SSC bei RT). Die stringenten Hybridisierungsbedingungen für die in III.6. beschriebenen Experimente waren 5 x SSC, 50% Formamid, 42°C, wobei der letzte Waschschritt nach der Hybridisierung in 0,1 x SSC bei 65°C stattfand.

8. In situ Hybridisierung

Die Herstellung der Quetschpräparate von Riesenchromosomen sowie die Präparatevorbehandlung erfolgte nach Pardue & Gall (1975).
Die Hybridisierungsproben wurden durch Nicktranslation (s.u.) mit ^{125}J-markiertem dCTP radioaktiv markiert. Die Hybridisierungslösung enthielt 1 x DH, 100µg/ml E.coli DNA, 20 mM Tris-HCl pH 7,2, 2 x SSC sowie die Probe in einer Konzentration von 0,3µg/ml. Hybridisiert wurde bei 55°C über Nacht. Gewaschen wurde 5 x in 2 x SSC bei 50°C. Die Hybridisierung wurde autoradiographisch nachgewiesen, wobei der sogenannte ´stripping´ Film von Kodak benutzt wurde.

9. Enzymreaktionen

Restriktionsenzyme: Für die Restriktionsverdaus wurden drei verschiedene Puffer verwendet, die sich im wesentlichen nur in ihrem Salzgehalt unterscheiden. (high: 10 mM Tris-HCl pH 7, 5, 10 mM $MgCl_2$, 100 mM NaCl; medium: 10 mM Tris-HCl pH 7,5, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT; low: 10 mM Tris-HCl pH 7,5, 10 mM $MgCl_2$, 1 mM DTT). Für die einzelnen Enzyme

wurden dann folgende Puffer verwendet: Alu I - medium, Ava I - medium, Bam HI - medium, Bgl I - medium, Bgl II - low, Cla I - low, Eco RI - high, Hae II - low, Hae III - low, Hha I - medium, Hind III - medium, Hinf I - medium, Hpa I - low, Hpa II - low, Pst I - medium, Pvu II - medium, Sac I - low, Sal I - high, Sau 3A - medium, Xba I - low.

Ligasereaktionen wurden durchgeführt in 50 mM Tris-HCl pH 7,2, 10 mM $MgCl_2$, 100 mM NaCl, 5 mM DTT, 1 mM ATP. Die DNA-Konzentration betrug zwischen 5-50µg/ml (Berechnung nach Dugaiczyk et al., 1975). Bei Ligierung von überstehenden Enden wurden ca. 10 U/ml Ligase eingesetzt, bei glatten Enden ca. 1000 U/ml. Erstere Ligation wurde über Nacht bei 15°C, zweitere über Nacht bei RT durchgeführt.

Die Nicktranslation mit Kornbergpolymerase wurde im wesentlichen nach den von Rigby et al. (1977) angegebenen Bedingungen durchgeführt. Es wurden in der Regel mindestens zwei verschiedene ^{32}P-markierte Nukleotide eingesetzt, so daß immer beide Stränge der simplen Sequenzen markiert wurden. Nicht eingebaute radioaktive Nukleotide wurden in der Regel auf einer kleinen Sephadex G 100 Säule (2, 5 ml in Pasteurpipette) mit TE als Säulenpuffer abgetrennt.

Die simple Sequenz GA/CT wurde aus ihren Dekamerprimern ($p(dAdG)_5$ und $p(dCdT)_5$) nach den von Byrd et al. (1965) angegebenen Bedingungen synthetisiert. Die Ausbeute dieser Synthese war sehr gering (ca. 1 ng DNA aus etwa 1µg Primern), da die radioaktiven Nukleotide nicht in der eigentlich notwendigen Konzentration eingesetzt werden konnten. Da bei dieser Synthese aber ein Produkt mit maximaler spezifischer

Aktivität entsteht, waren die Hybridisierungsexperimente damit dennoch zufriedenstellend.

Klenowpolymerase: Die Klenowpolymerase ist ein Teilfragment der Kornbergpolymerase, das keine 5′→3′ Exonukleaseaktivität mehr besitzt (Klenow & Henningsen, 1970). Dieses Enzym wurde eingesetzt, um 5′-überhängende Enden aufzufüllen, bzw. sie radioaktiv zu markieren (zur Sequenzierung). Da das Enzym aber noch die 3′→5′ Exonukleaseaktivität besitzt, wurde es auch dazu verwendet, 3′-überhängende Enden zurückzuschneiden, bzw. zu markieren. Die Enzymreaktion wurde in medium-Puffer (s.o.) und für 30 min bei RT durchgeführt. Markierte und nicht markierte Deoxynukleotide wurden nach den Erfordernissen des entsprechenden Endes eingesetzt, wobei in jedem Fall immer das Nukleotid anwesend war, an dem die 3′→5′ Exonukleaseaktivität stoppen sollte (die Polymeraseaktivität ist ca. 1000 mal größer!).

Polynukleotidkinase: Mit diesem Enzym wurde DNA nach Phosphataseverdau zur Sequenzierung 5′ markiert, wobei nach den von Maxam & Gilbert (1980) angegebenen Bedingungen verfahren wurde. RNA wurde nach partieller Hydrolyse 5′-markiert, unter den von Maizels (1976) angegebenen Bedingungen.

10. Klonierung

Alle verwendeten Standardklonierungstechniken (Medien, Transformation, in vitro packaging von Phagen, Phageninfektion, Plasmid- und Phagenisolierung, Klonbankscreening) sind in Maniatis et al. (1982) beschrieben. Die

Mikroklonierungstechniken sind in Scalenghe et al. (1981) beschrieben.

11. Sequenzierung

Es wurde die von Maxam & Gilbert (1980) beschriebene chemische Degradationsprozedur verwendet. Die DNA-Fragmente wurden entweder 5´-markiert (Polynukleotidkinase, s.o.) oder 3´-markiert (mit Klenowpolymerase, s.o.). Die nachgeschnittenen Fragmente wurden auf Agarosegelen aufgetrennt und mit Hilfe einer modifizierten freeze-squeeze Technik (Tautz & Renz, 1983) isoliert. Es wurden hochauflösende, ultradünne Sequenzierungsgele verwendet (Garoff & Ansorge, 1981).

ERGEBNISSE

I. Simple Sequenzen in DNA

I.1. Nachweis durch dot-blot Hybridisierungen

Mit einem einfachen, qualitativen Hybridisierungsexperiment wurde getestet, ob und welche simplen Sequenzen in der DNA von Eukaryonten zu finden sind.

Getestet wurden sechs phylogenetisch unterschiedliche Organismen, nämlich Mensch, Xenopus, Drosophila, Seeigel, Weizen und Hefe. Jeweils 1µg der isolierten DNA von diesen Organismen wurde auf Nitrocellulose fixiert und gegen radioaktiv markierte Proben von synthetischen simplen Sequenzen, nämlich AA/TT,GG/CC,AT/TA,GC/CG,GT/CA und GA/CT hybridisiert. Die Hybridisierungsbedingungen wurden dabei so kalkuliert, daß auch kurze Sequenzen (>15 bp) noch ein Signal ergeben sollten (vgl. Methoden).

Es zeigte sich, daß alle diese simplen Sequenzen mit allen genomischen DNAs hybridisieren, mit Ausnahme von GC/CG in Weizen und Hefe (siehe Abb. 2A).

Die Hybridisierungsresultate mit GC/CG, aber auch mit AT/TA sind andererseits wahrscheinlich nicht sehr zuverlässig, da diese Proben selbstkomplementär sind und nach der Denaturierung sofort wieder in einer intramolekularen Reaktion renaturieren. Es ist daher erstaunlich, daß überhaupt eine Hybridisierung an eine andere DNA gefunden werden kann, aber die gezeigten Resultate sind reproduzierbar.

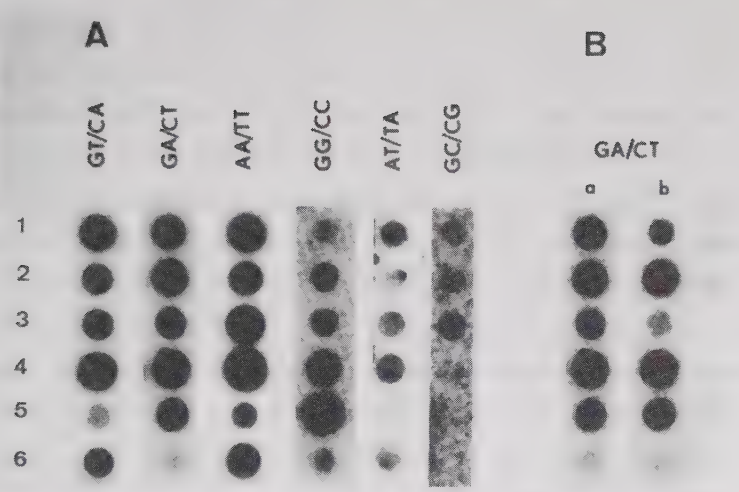


Abb. 2: Dotblothybridisierungen der verschiedenen simplen Polymere an die DNA von sechs verschiedenen Eukaryonten

Je 1 µg DNA von: 1 = Mensch, 2 = Xenopus, 3 = Drosophila, 4 = Seeigel, 5 = Weizen, 6 = Hefe.
A: Hybridisierung unter nicht stringenten Bedingungen.
B: Hybridisierung nur mit GA/CT, Vergleich zwischen nicht stringentem (a) und stringentem Waschschritt (b).

Alle sechs Polymere hybridisierten unter den gleichen Bedingungen nicht gegen DNA von E.coli oder λ (WT) DNA.

Da die für diese Experimente gewählten Hybridisierungsbedingungen auch einen hohen Falschpaarungsanteil zulassen (theoretisch bis 50% bei sehr langen Sequenzen, vgl. IV. 1.), wurden die Experimente noch einmal wiederholt, aber mit einem stringenten Waschschritt nach der Hybridisierung (siehe Methoden). Vergleicht man die Resultate von diesen Experi-

menten mit den ersteren, so erkennt man nur wenige signifikante Unterschiede. Neben einer allgemeinen Reduzierung der Signale (besonders in der Hybridisierung mit AT/TA und GC/CG) ergibt sich nur noch in der Hybridisierung mit GA/CT im Mensch und in Drosophila ein deutlicher Unterschied (Abb. 2B). In Drosophila melanogaster gibt es einen Satelliten, der aus dem Pentanukleotid AAGAG besteht (Sederoff et al., 1975). Es ist wahrscheinlich, daß eine Kreuzhybridisierung zwischen GA/CT und diesem Satelliten stattfindet. Da dabei aber viele Fehlpaarungen entstehen, schmelzen die Hybride unter den Bedingungen des zweiten Experiments wieder ab. Eine ähnliche Erklärung ist auch für die menschliche DNA denkbar.

In allen anderen Fällen blieb die relative Höhe des Signals gleich. Dies deutet darauf hin, daß tatsächlich einfache Sequenzen in Eukaryonten als ununterbrochene Stücke von mindestens 20 bp Länge vorkommen und zwar in mehr, als statistischen Mengen. Wenn man eine Gleichverteilung aller Nukleotide annimmt, kann man die Wahrscheinlichkeit w, daß eine bestimmte Sequenz zufällig vorkommt, nach der einfachen Formel $w = 1/4^n$ berechnen, wobei n die Länge der Sequenz in bp ist. Man erhält dann für n = 15: $w = 1 \times 10^{-9}$ und für n = 20: $w = 1 \times 10^{-12}$. Eine simple Sequenz mit 20 bp Länge sollte daher in eukaryontischen Genomen nicht mehr auf Grund von zufälliger Anordnung der Basen entstehen, da deren Genomgröße in der Regel nicht mehr als 10^{10} bp beträgt. Aber auch kürzere Sequenzen sollten höchstens ein, oder wenige Male zufällig vorkommen.

Alle weiteren Experimente wurden im wesentlichen nur noch mit den simplen Sequenzen GT/CA, GA/CT und AA/TT durchgeführt.

I.2. Quantifizierung

Mit ähnlichen, wie den oben gezeigten Experimenten lassen sich auch quantitative Aussagen machen. Dazu vergleicht man die Hybridisierungssignale von den genomischen DNAs, mit dem Signal von einer klonierten DNA mit einem bekannten Gehalt der einfachen Sequenz (vgl. Abb. 3). Für diese Versuche wurden einige Klone verwendet, die später beschrieben werden. Die Hybridisierungen wurden unter den gleichen Bedingungen wie in I.1. durchgeführt, allerdings mit dem stringentem Waschschritt (vgl. Methoden).

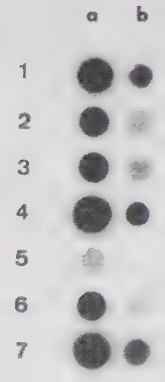


Abb. 3: Beispiel für die Quantifizierung der Hybridisierung

Hybridisierung an dot blots mit GT/CA. 1-6 wie in Abb. 2, 7 = 0,2 µg linearisiertes Plasmid pDv-182 (a); jeweils 1:10 Verdünnungen (b); jeweils 1:100 Verdünnungen (c).

In der Tabelle I sind die Ergebnisse aus diesen Versuchen für die einfachen Sequenzen GT/CA, GA/CT und AA/TT zusammengefaßt. Die verwendete Methode ist relativ grob und Fehler im Rahmen eines Faktors zwei bis drei sind denkbar, vor allem, weil nichts über die Längenverteilung der einfachen Sequenzen in den Genomen bekannt ist. Sehr kurze (< 20bp) Sequenzen hybridisieren nämlich wahrscheinlich relativ zu schwach (vgl. IV.1.).

Andererseits ergeben wahrscheinlich Sequenzen, die wesentlich länger als die Sequenz in dem entsprechenden Vergleichsklon sind, relativ zu starke Signale. Eine Korrektur für diese Fehler wurde nicht versucht.

	Genom Größe	GT/CA		GA/CT		AA/TT	
	bp	%	bp	%	bp	%	bp
Mensch	$3{,}3 \times 10^9$	0,5	2×10^7	0,1	3×10^6	0,3	1×10^7
Xenopus	$3{,}1 \times 10^9$	0,2	6×10^6	0,3	9×10^6	0,05	2×10^6
Drosophila	$1{,}4 \times 10^8$	0,2	3×10^5	0,02	3×10^4	0,2	3×10^5
Seeigel	$8{,}6 \times 10^8$	0,8	7×10^6	0,5	4×10^6	0,8	7×10^6
Weizen	$1{,}7 \times 10^{10}$	0,03	5×10^6	0,04	7×10^6	0,005	8×10^5
Hefe	$4{,}0 \times 10^7$	0,1	4×10^4	0,0002	80	0,1	4×10^4

Tab. 1: Abschätzung des relativen Anteils von drei simplen Sequenzen in sechs verschiedenen Eukaryonten.

In der Tabelle wird jeweils der relative Anteil (%) und der absolute Anteil (bp) der einzelnen simplen Sequenzen in den Genomen angegeben. Die Zahlen für die Genomgröße sind entnommen aus (Lewin, 1980).

In den meisten Fällen wurden Werte zwischen 0,1 bis 0, 5% gefunden, also ein vergleichsweise ähnlicher Anteil bei den drei verschiedenen Sequenzen.

Insgesamt machen simple Sequenzen also einen substantiellen Anteil in eukaryontischen Genomen aus. Es ist leicht denkbar, daß 2-5% der eukaryontischen DNA aus den verschiedenen simplen Sequenzen bestehen (Satelliten noch nicht mit eingerechnet).

I.3. Genomblothybridisierungen

Die nächsten Experimente sollten Aufschluß über die Verteilung der simplen Sequenzen in der DNA geben. Dazu wurden die genomischen DNAs der verschiedenen Spezies mit dem Restriktionsenzym Eco RI verdaut, auf einem Agarosegel nach Größe aufgetrennt, auf Nitrocellulosefilter geblottet und gegen die simplen Sequenzen hybridisiert, wobei der stringente Waschschritt (s.o.) mit eingeschlossen wurde.

Wenn die einfachen Sequenzen, ähnlich wie die Satelliten DNA sehr lange, ununterbrochene Stücke bilden, dann sollte man bei diesem Experiment nur wenige hochmolekulare Banden sehen. Sind sie dagegen über das ganze Genom verteilt, so würde man sehr viele hybridisierende Banden erwarten, die nicht mehr einzeln auflösbar sind und daher einen zusammenhängenden Schmier bilden. Letzteres wurde im wesentlichen auch gefunden. Man sieht im allgemeinen einen Schmier mit wenigen Banden (Abb. 4). Nur in der Hefe sieht man hauptsächlich Banden, im Fall von GA/CT sogar nur eine. Da das Hefegenom nur eine geringe Komplexität (ca. 4×10^7 bp) besitzt, konnte dies auch erwartet werden. Auch in der Hybridisierung mit GA/CT in Drosophila zeichnet sich aus dem gleichen Grund bereits ein Bandenmuster ab.
Die anderen sichtbaren Banden deuten entweder auf eine Clusterung der entsprechenden simplen Sequenz auf einem Fragment hin, oder sie kommen dadurch zustande, daß eine andere repetetive Sequenz, die die Bande bildet, eine simple Sequenz enthält.

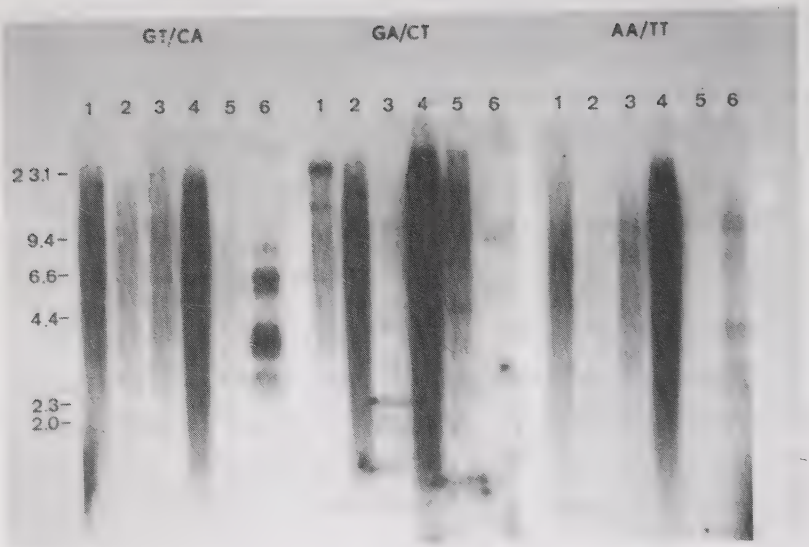


Abb. 4: Hybridisierung der simplen Polymere an Genomblots von sechs verschiedenen Eukaryonten.

Es wurde jeweils 1µg Eco RI verdaute DNA pro Spur aufgetragen. 1 = Mensch; 2 = Xenopus; 3 = Drosophila; 4 = Seeigel; 5 = Weizen; 6 = Hefe.
Bei längerer Exposition der Filter sieht man auch in den Spuren 5 (beim 1. und 3. Filter) einen homogenen Schmier.

I.4. In situ Hybridisierungen

Ein anderes Experiment, das Aufschluß über die Verteilung der simplen Sequenzen im Genom gibt, ist die in situ Hybridisierung an polytäne Chromosomen. Es ist bekannt, daß Satelliten DNAs hauptsächlich in den Centromeren dieser Chromosomen lokalisiert sind (Lewin, 1980, Kap 19). Die Hybridisierung mit GA/CT und GT/CA zeigt jedoch, daß die Chromosomen in ihrer ganzen Länge markiert werden (Abb. 5). Zumindest bei der GA/CT-Hybridisierung (Abb. 5, linker Teil) ist deutlich, daß die Centromeren sogar weniger hybridisieren. Über den Nukleoli sieht man bei beiden Experimenten nahezu keine Markierung. Einige Teile der Chromosomen sehen so aus, als wären sie dichter markiert, als andere. Dies könnte darauf hindeuten, daß die simplen Sequenzen in einigen Chromosomenbereichen geclustert sind. Diese Ansicht wird auch durch die späteren Versuche (s. III.1. und III.4.) unterstützt. Andererseits könnte dieser Befund aber auch teilweise durch einen unterschiedlichen Spreitungsgrad der Chromosomen zustande kommen.

Da die Hybridisierungsproben mit ^{125}J markiert waren, erhält man keine genügend große Auflösung, um zu entscheiden, ob Banden oder Interbanden stärker hybridisieren.
Versuche, bei denen die Hybridisierung durch, mit einem Fluoreszenzfarbstoff markierte Antikörper nachgewiesen wird (Langer et al., 1981), wurden auch durchgeführt. Es zeigte sich jedoch, daß diese Methode noch nicht sensitiv genug ist, um die Hybridisierung an so kurze Stücke (< 150bp), wie sie die simplen Sequenzen bilden (vgl. III.5.), nachzuweisen.

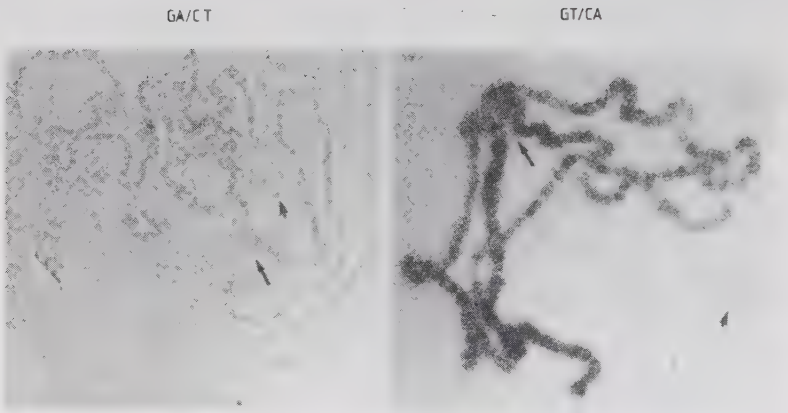


Abb. 5: In situ Hybridisierung von GA/CT und GT/CA an polytäne Chromosomen von D. virilis

In beiden Bildern sind jeweils die Nukleoli (kurzer Pfeil) sowie die Centromere (langer Pfeil) gekennzeichnet. Der Nukleolus im rechten Bild ist nicht ganz im Focus. Beide Hybrisisierungsproben hatten die gleiche spezifische Aktivität (ca. 5 x 10^7 cpm/µg). Statt synthetischem GA/CT wurde jedoch das Plasmid pMm-2 (vgl. III.4.5.) benutzt. Expositionszeit 24h.

II. Simple Sequenzen in RNA

Viele mittelrepetetive Sequenzen wurden auch in der RNA gefunden (Davidson & Posakony, 1982). In den folgenden Versuchen wurde daher getestet, ob dies auch für die simplen Sequenzen zutrifft.

II.1. Hybridisierung an Gesamt-RNA

Von den sechs, in dieser Arbeit verwendeten Eukaryonten wurde die Gesamt-RNA isoliert und auf Formaldehyd-Agarose Gelen aufgetrennt. Von diesen Gelen kann die RNA direkt auf Nitrocellulose geblottet werden. Die Filter wurden dann mit drei verschiedenen simplen Sequenzen hybridisiert, nämlich mit GT/CA, GA/CT und poly-rA. Letzteres wurde statt AA/TT verwendet, um Hybridisierung gegen die poly-rA Stücke der mRNA zu vermeiden. (Die verwendete poly-rA Präparation hybridisierte im Kontrollexperiment nicht gegen, an Nitrocellulose fixiertes poly-rA, enthielt also kein poly-U.) Das Ergebnis dieser Hybridisierungen zeigt Abb. 6. Alle drei simplen Sequenzen hybridisierten gegen die RNA von allen getesteten Eukaryonten. Besonders stark hybridisierte poly-rA gegen die RNA von Seeigel. Es ist allerdings auch bereits bekannt, daß diese RNA viele oligo-U Sequenzen enthält (Duncan & Humphreys, 1983). Alle Hybridisierungen zeigen im wesentlichen einen Schmier mit vereinzelten Banden. Der Schmier ist wahrscheinlich auf eine Vielzahl von Transkripten verschiedener Länge zurückzuführen und nicht auf eine Degradation der RNA. Die Filter wurden nämlich nach der Hybridisierung noch gefärbt, wobei die ribosomalen Banden

keine wesentliche Degradation zeigten (mit Ausnahme der RNA von Xenopus, Spur 2).

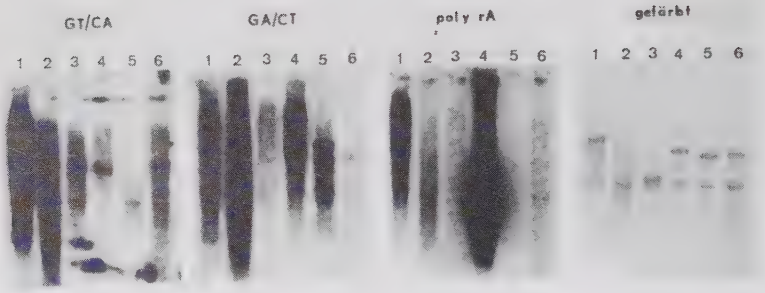


Abb. 6: Hybridisierung der simplen Polymere an Gesamt-RNA von sechs verschiedenen Eukaryonten.

Es wurden jeweils 5µg Gesamt-RNA pro Spur aufgetragen. Der rechte Filter wurde gefärbt. 1 = Mensch, RNA aus Fibroblastenzellen; 2 = Xenopus, RNA aus der Leber; 3 = Drosophila, RNA aus Kc-Zellen; 4 = Seeigel, RNA aus unbefruchteten Eiern; 5 = Weizen, RNA aus grünen Blättern; 6 = Hefe, RNA aus Zellen aus der frühen log-Phase.

II.2. Hybridisierung an fraktionierte Kc-Zell RNA

Um die Verteilung der simplen Sequenzen in der Zelle zu untersuchen, wurde eine Zellfraktionierung von Drosophila Kc-Zellen vorgenommen. Zum einen wurde die RNA in eine Kern- und eine Cytoplasmafraktion unterteilt. Diese wurden auf Formaldehyd-Agarosegelen aufgetrennt, geblottet und gegen die

simplen Sequenzen hybridisiert (Abb. 7, Spuren a+b). Es zeigte sich, daß die Sequenz GT/CA im Kern wie im Cytoplasma gleichmäßig vertreten ist. Zieht man allerdings in Betracht, daß die Kernfraktion nicht ganz cytoplasmafrei war, (noch ein starker Anteil ribosomaler RNA, vgl. gefärbten Filter in Abb. 7) kommt man zu dem Schluß, daß GT/CA im Cytoplasma sogar stärker vertreten ist.

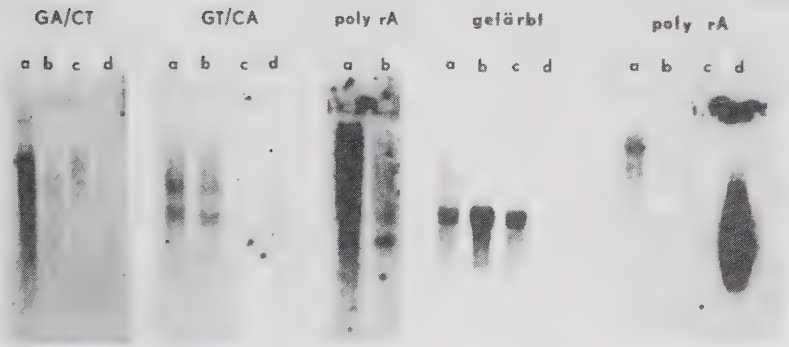


Abb. 7: Hybridisierung der simplen Polymere an RNA-Fraktionen aus Drosophila Kc-Zellen.

Es wurden jeweils equivalente Mengen der unterschiedlichen Fraktionen aufgetragen. a = Kern-RNA (ca. 5µg); b = cytoplasmatische RNA (ca. 15µg); c = poly-A^- RNA (ca. 5µg); d = poly-A^+ RNA (ca. 0, 3µg). Der äußerste rechte Filter entspricht dem dritten Filter, aber mit einer kürzeren Expositionszeit.

Die simple Sequenz GA/CT hybridisiert allerdings deutlich stärker mit der Kernfraktion (Abb. 7). Bei der Hybridisierung mit poly-rA sieht man sogar im wesentlichen nur eine Bande in der cytoplasmatischen RNA. Interessanterweise findet man an dieser Stelle auch auf dem gefärbten Filter (s. Abb. 7) eine Bande. Es muß sich also um eine RNA-Spezies handeln, die in sehr großen Mengen transkribiert wird.

Als zweites wurde die Gesamt-RNA aus Kc-Zellen noch in die poly-A^+ und poly-A^- RNA fraktioniert. Die Hybridisierung gegen diese Fraktionen zeigte auch wieder ein unterschiedliches Verhalten der simplen Polymere. GT/CA wird hauptsächlich in der poly-A^+ Fraktion gefunden, während GA/CT hauptsächlich in der poly-A^- Fraktion vorkommt (Abb. 7, Spuren c+d).
Die Hybridisierung mit poly-rA führte in diesem Experiment dagegen nur zu nicht interpretierbaren Ergebnissen. Man findet nämlich in der poly-A^+ Fraktion eine sehr starke Hybridisierung, wesentlich stärker, als es von der Menge der aufgetragenen RNA her zu erwarten wäre (Abb. 7, rechter Filter). Wahrscheinlich handelt es sich dabei um ein Artefakt, das bei der Fraktionierung auf der Oligo-dT-Cellulose Säule entsteht. Möglicherweise werden dabei fines ausgewaschen, die auch im Gel wandern und dann auf Grund ihres oligo-dT Anteiles ein Hybridisierungssignal ergeben. Dieser Effekt wurde allerdings auch erhalten, wenn die Säule sehr sorgfältig und lange vorgewaschen wurde.

II.3. Hybridisierung an stadien- und gewebespezifische Drosophila RNA

Um zu testen, ob es organspezifische Transkripte gibt, die simple Sequenzen enthalten, wurde aus drei Organen von D.virilis Larven in einem zu diesem Zweck entwickelten Mikroverfahren (siehe Methoden) RNA isoliert. Diese RNA wurde wieder auf Gelen aufgetrennt, geblottet und in diesem Fall nur gegen GT/CA hybridisiert. Es zeigte sich, daß das Hybridisierungsmuster in den drei Organen tatsächlich unterschiedlich ist (Abb. 8).

Die Hybridisierung wurde in diesem Fall allerdings nicht mit synthetischem GT/CA vorgenommen, sondern mit einem aus dem Klon pDv-142 (vgl. III.4.2.) subkloniertem Fragment, das drei GT/CA-Stücke trägt (das 0, 55 kb Hae II Fragment, vgl. Restriktionskarte). Es ist daher nicht ausgeschlossen, daß noch andere, nicht simple Sequenzen an dieser Hybridisierung beteiligt waren.

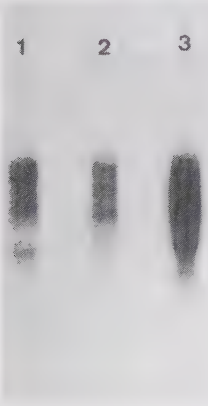


Abb. 8: Hybridisierung von GT/CA an Gesamt-RNA aus verschiedenen Organen von D.virilis Larven.

1 = Drüsen
2 = Fettköper
3 = Gehirn

Ein weiterer Versuch sollte zeigen, ob simple Sequenztranskripte in der Entwicklung reguliert sind. Dazu wurde RNA von verschiedenen Entwicklungsstadien von D.melanogaster gegen simple Sequenzen hybridisiert.
Es zeigte sich, daß das Hybridisierungsmuster im wesentlichen während der Entwicklung gleich bleibt (Abb. 9), der relative Anteil sich allerdings verändert. Eine quantitative Auswertung dieses Versuchs ist allerdings nicht möglich, da die aufgetragene RNA-Menge in den einzelnen Spuren nicht gleich ist, was auf dem gefärbten Filter zu sehen ist. Die Unterschiede in den Hybridisierungssignalen sind jedoch stärker. Insbesondere das späte Larvenstadium (Spuren 4) sowie das späte Puppenstadium (Spuren 2) hybridisieren jeweils wesentlich schwächer. Dagegen scheinen während der Entwicklung zur Puppe, also beim Übergang vom 3. Larvenstadium zum frühen Puppenstadium (Spuren 4 und 3) noch einmal simple Sequenzen transkribiert zu werden, die jedoch rasch wieder degradiert werden (in der späten Puppe, Spuren 2).

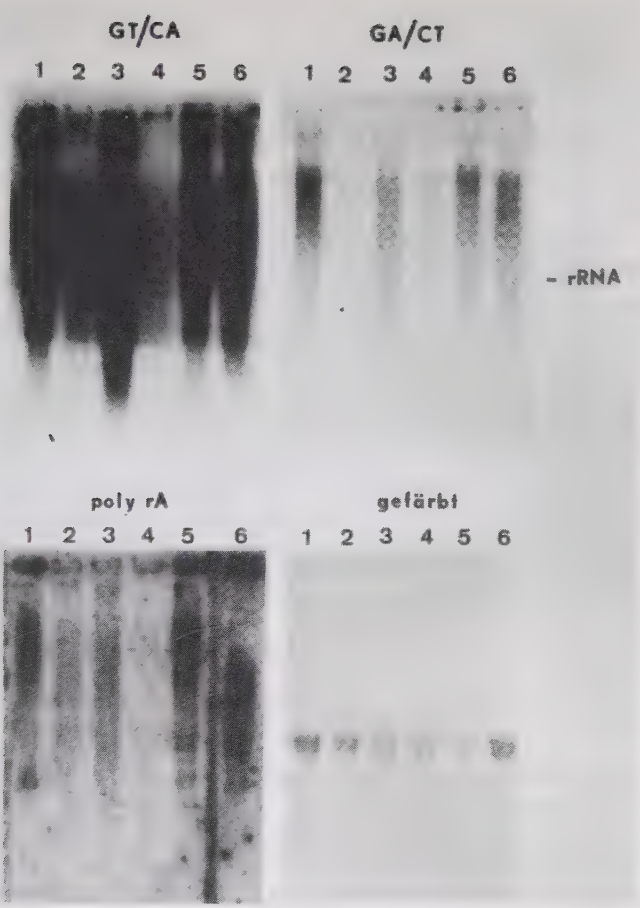


Abb. 9: Hybridisierung von drei simplen Sequenzen an die Gesamt-RNA aus verschiedenen Entwicklungsstadien von Drosophila melanogaster.

1 = Fliegen; 2 = späte Puppe; 3 = frühe Puppe; 4 = Larven des dritten Larvenstadiums; 5 = 12-21h alte Embryos; 6 = 3-5h alte Embryos. Die Position der rRNA ist bei den oberen Filtern markiert.

III. Klonierung und Sequenzierung

Um die simplen Sequenzen genauer zu charakterisieren, wurden genomische, klonierte DNA-Stücke aus Drosophila virilis Klonbänken, die mit einfachen Sequenzen hybridisierten, isoliert. Von einigen ausgewählten Klonen wurde dann eine genaue Restriktionskarte angefertigt und die Fragmente, die mit der einfachen Sequenz hybridisierten, wurden sequenziert. Drei Typen von Klonbänken wurden zu diesem Zweck herangezogen. Die meisten hier verwendeten Klone stammen aus zwei Mikroklonbänken. Die Mikroklonierung ist eine von Scalenghe et al. (1981) entwickelte Methode, bei der unter dem Mikroskop definierte Stücke von Chromosomenpräparaten isoliert werden. Die DNA aus diesen Stücken wird mit Eco RI geschnitten und in dem Insertionsvektor λ 641 (Murray et al., 1977) kloniert.

Für die vorliegende Arbeit wurden zwei Klonbänke aus zwei Puffs vom Ende des dritten Chromosoms bzw. vom Ende des X-Chromosoms (Abb. 10) von Drosophila virilis hergestellt (in Zusammenarbeit mit Jan-Eric Edström, EMBL, Heidelberg). Die Klone pDv-1 und pDv-141 stammten dagegen aus Klonbänken, bei denen die gesamte genomische DNA von D.virilis partiell mit Sau 3A verdaut, größenfraktioniert und in die Bam HI-Stelle des Phagen λ EMBL 2 bzw. des Plasmids pBR 322 kloniert wurde. Die erstere Klonbank wurde hergestellt von Manfred Renz (EMBL, Heidelberg), die zweitere von Isolde Riede (Riede, 1982) (MPI für Biologie, Tübingen).

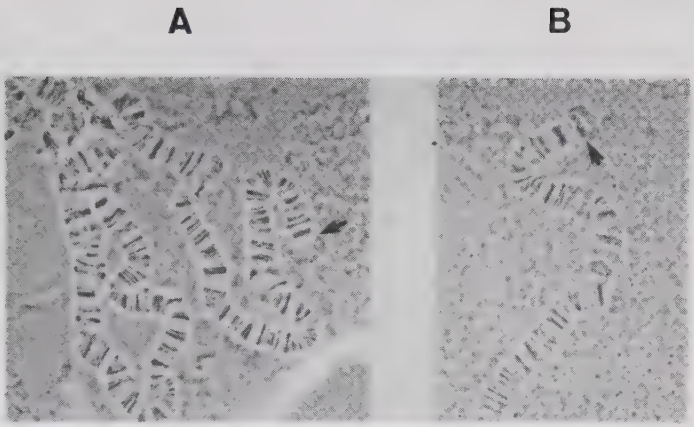


Abb. 10: Polytäne Chromosomen von D.virilis

Die Pfeile in den Bildern bezeichnen die Stellen, an denen die Chromosomen zur Mikroklonierung geschnitten wurden. A = drittes Chromosom; B = X-Chromosom

III.1. Isolierung und Charakterisierung von Klonen aus zwei Mikroklonbänken

Die Klonbank vom X-Chromosom bestand aus 108 Klonen, die vom dritten Chromosom aus 214. Alle diese Klone wurden in einer Benton/Davies Prozedur (Benton & Davies, 1977) auf Kreuzhybridisierung mit den simplen Sequenzen getestet. Zusätzlich wurde noch poly-A^+ RNA aus den Speicheldrüsen des späten dritten Larvenstadiums isoliert, 5´ ^{32}P-markiert und ebenfalls gegen die oben genannten Klone hybridisiert. Das Ergebnis dieser Hybridisierung ist in Tabelle II zusammengefaßt. Obwohl die Klonbank vom X-Chromosom kleiner war (nur halb so viele Klone), enthielt sie wesentlich mehr Klone, die gegen simple Sequenzen hybridisierten. Auch dieses Ergebnis deutet auf eine Clusterung von simplen Sequenzen in bestimmten Chromosomenbereichen hin. Viele der Klone, die mit der RNA hybridisierten, taten dies auch mit einer der simplen Sequenzen GT/CA oder GA/CT. (Die Bedingungen für die RNA-Hybridisierung waren so gewählt, daß keine Hybridisierung an kurze (< 50 bp) AA/TT-Stücke stattfinden konnte).

Es ist wahrscheinlich, daß weitere Klone nur auf Grund ihres Gehalts einer anderen simplen Sequenz mit der RNA hybridisierten.

Zehn Klone wurden etwas näher analysiert. Dazu wurde DNA von ihnen isoliert, mit Eco RI geschnitten und auf einem Agarosegel aufgetrennt. Man kann auf diese Art die klonierten Fragmente von der λ-DNA, die nur zwei Fragmente mit 7,9 und 31,4 kb liefert, trennen. Alle Klone, bis auf einen (Klon X-66), enthielten nur ein zusätzliches Eco RI

Klon	GT/CA	CA/CT	AA/TT	RNA
X - 7	x			(x)
X - 12		x		(x)
X - 15			x	
X - 21				x
X - 23			x	
X - 24				x
X - 27		x	x	(x)
X - 28		(x)		
X - 29			x	
X - 35				(x)
X - 38	x			(x)
X - 40		x	x	x
X - 43	x			x
X - 49		(x)		
X - 50			x	x
X - 51				x
X - 52			x	
X - 54	x			(x)
X - 57	x		x	x
X - 59				x
X - 61			x	x
X - 63	x			x
X - 64		x	x	(x)
X - 65			x	x
X - 66	(x)			
X - 67	x			(x)
X - 68			x	
X - 69				(x)
X - 70	x			(x)
X - 74		(x)	x	
X - 83			x	
X - 86			x	
X - 91		(x)	x	
X - 95			x	
X - 96				x
X - 99			x	x
X - 103	x		x	x
X - 104				x
X - 107	x			(x)
Summe:	11	8	19	26
3 - 19		(x)		
3 - 30				x
3 - 63				x
3 - 112				x
3 - 120				(x)
3 - 161	x			x
3 - 163	x			x
3 - 165		x		x
3 - 193				(x)
Summe:	2	2	0	8

Tab. II: Auflistung der Klone aus den Mikroklonbänken, die mit einer simplen Sequenz oder mit RNA hybridisieren.

Die x-Zeichen geben an, welche Probe jeweils mit dem Klon hybridisierte. In der Tabelle wird im wesentlichen nicht zwischen starker und schwacher Hybridisierung unterschieden. Lediglich bei sehr schwachen Signalen ist das x-Zeichen eingeklammert.

Fragment (Abb. 11A). Bei Klon X-66 ist nicht klar, ob die enthaltenen Fragmente Stücke repräsentieren, die auch im Genom zusammenliegen, oder ob sie durch Religierung unterschiedlicher Stücke entstanden sind. Aus den Erfahrungen mit der Mikroklonierung ist aber der erste Fall wahrscheinlicher.

Die Klone X-53 und X-103 enthalten gleich große klonierte Fragmente und zeigen auch ähnliche Hybridisierungseigenschaften (s.u.), so daß vermutet werden muß, daß sie identisch sind.

Das oben genannte Gel wurde auf Nitrocellulose geblottet und nacheinander gegen drei verschiedene Proben hybridisiert. Die erste Probe war ^{32}P-markierte poly-A$^+$ RNA aus Drüsen des späten dritten Larvenstadiums von D.virilis. Alle Klone hybridisierten mehr oder weniger stark mit dieser Probe (Abb. 11B). Danach wurde die gleiche Hybridisierung durchgeführt, aber mit einem hohen Überschuß an nicht markiertem GT/CA (Abb. 11C). Unter diesen Bedingungen wird die Hybridisierung der RNA an GT/CA völlig unterdrückt (vgl. Abb. 11B mit 11C). Nur noch zwei Klone (X-43 und X-104) hybridisieren stark, drei andere sehr schwach (X-57, X-103 und 3-165). Schließlich wurde der Filter noch mit radioaktivem GT/CA hybridisiert. Bis auf zwei Klone (X-104 und 3-165) zeigten nun alle ein starkes Signal (Abb. 11D). Klon 3-165 enthält allerdings eine andere simple Sequenz, nämlich GA/CT (vgl. Tab. II). Auch die beiden anderen, unter Kompetetionsbedingungen noch schwach hybridisierenden Klone X-103 und X-57 enthalten noch andere simple Sequenzen

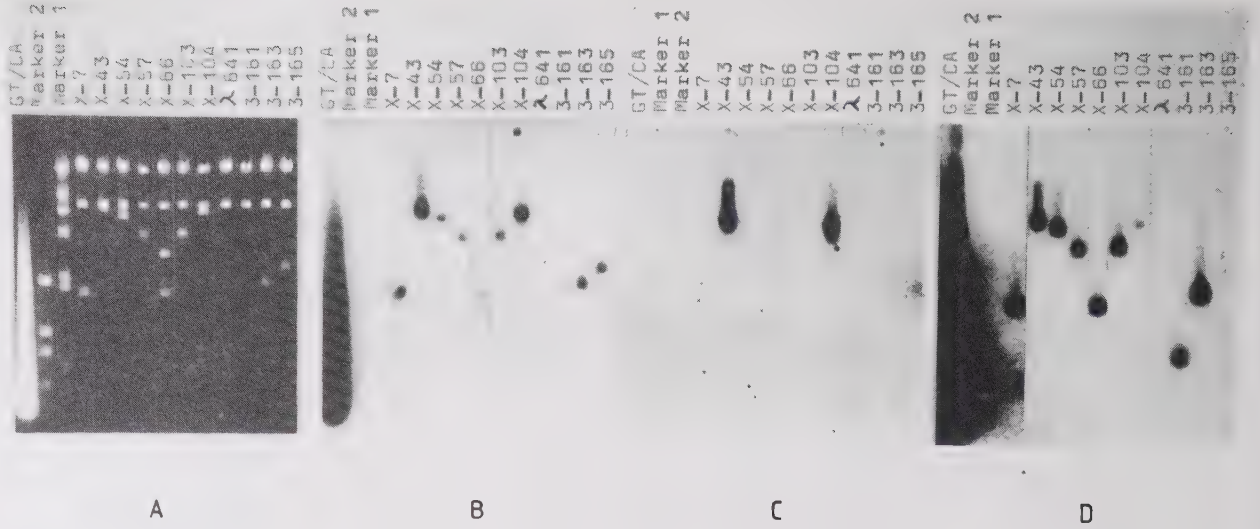


Abb. 11: Hybridisierung an Inserts von Klonen aus der Mikroklonbank mit verschiedenen Proben

Die λ-DNAs wurden jeweils mit Eco RI geschnitten und auf einem 0,8% Agarosegel aufgetrennt. A = gefärbtes Gel; B = Hybridisierung mit RNA aus Drüsen des späten dritten Larvenstadiums von D.virilis; C = Hybridisierung wie in B, aber unter Kompetetion von 1µg/ml nicht markiertem GT/CA; D = Hybridisierung mit radioaktivem GT/CA.

(vgl. III.4.3., Klon pDv-161). Es bleibt also nur ein Klon, der gegen RNA hybridisiert, in dem aber noch keine simple Sequenz gefunden wurde. Die Auswertung aller dieser Experimente ist in Tabelle III noch einmal zusammengefaßt.

Klon	RNA	RNA + GT/CA Komp.	GT/CA	Insertgröße
X - 7	x	-	xxx	1,9 kb
X - 43	xxx	xxx	xxx	7,3 kb
X - 54	x	-	xxx	6,1 kb
X - 57	x	(x)	xxx	4,5 kb
X - 66	x	-	xxx	3,3+1,9+0,6 kb
X - 103	x	(x)	xxx	4,5 kb
X - 104	xxx	xxx	(x)	6,7 kb
3 - 161	(x)	-	xx	0,85 kb
3 - 163	x	-	xxx	2,2 kb
3 - 165	x	x	-	2,8 kb

Tab. III: Charakterisierung von zehn Klonen aus der Mikroklonbank. Die Tabelle ist eine Zusammenfassung der in Abb. 11 gezeigten Hybridisierungen. Die Stärke der Hybridisierungssignale wurde mit der Zahl der x-Zeichen angedeutet.

III.2. Klone, die für die Sequenzierung ausgewählt wurden

Im folgenden werden die zur Sequenzierung ausgewählten Klone aufgelistet. Es wird jeweils die systematische Bezeichnung (die später eingeführt wurde) vorangestellt und kurz beschrieben, woher sie stammen bzw. wie sie konstruiert wurden.

Klon pDv-1

Der Klon ist ein, in die Eco RI-Stelle von pBR 322 kloniertes, 6,8 kb großes Eco RI Fragment. Dieses Fragment wurde aus einem λ-Klon (λ2, 33 kb großes Insert) isoliert, der aus der genomischen D.virilis Klonbank in EMBL 2 stammt (vgl. III.).

Klon pDv-142

Der Klon ist ein subkloniertes Stück aus einem Klon aus der genomischen D.virilis Klonbank in pBR 322 (vgl. III.) (Klon B 11, 20, 5 kb großes Insert). Das Stück ist flankiert von einer Hind III und der Sal I Stelle aus dem Vektor und ist 6 kb groß. Es wurde zwischen die Hind III/Sal I Stellen von pBR 322 kloniert (vgl. auch III.4.2.).

Klon pDv-161

Dieser Klon enthält das, in die Eco RI-Stelle von pBR 322 klonierte, 4, 45 kb große Eco RI-Fragment aus dem λ-Klon X-103 aus der Mikroklonbank vom X-Chromosom. Dieser hybridisierte im Vorscreening mit AA/TT und GT/CA (vgl. Tab. II).

Klon pDv-182

Dieser Klon enthält ein 0, 2 kb großes, zwischen die Eco RI/Hind III Stellen des Plasmids pUC 9 kloniertes Fragment. Dieses stammt aus dem 0,8 kb großen Insert des λ-Klons 3-161 aus der Mikroklonbank vom 3. Chromosom und hybridisierte im Vorscreening mit GT/CA (vgl. Tab. II).

Klon pDv-191

Dieser Klon enthält das, in die Eco RI-Stelle von pBR 322 klonierte, 2,8 kb große Eco RI-Fragment aus dem λ-Klon 3-165 aus der Mikroklonbank vom dritten Chromosom. Dieser hybridisierte im Vorscreening mit GA/CT (vgl. Tab. II).

III.3. Strategie für die Restriktionskartierung

Von den oben genannten Klonen wurden Restriktionkarten angefertigt und die Fragmente, die die jeweilige Kreuzhybridisierung zeigten, wurden möglichst genau lokalisiert. Dadurch war es dann möglich, die entsprechenden Bereiche gezielt zu sequenzieren. Im folgenden sind an einigen Beispielen die zur Restriktionskartierung verwendeten Strategien und Techniken beschrieben.

Der erste Schritt war in der Regel ein Vergleich zwischen einem Verdau des Klons mit dem fraglichen Enzym und einem Doppelverdau mit Eco RI und diesem Enzym. Da die Restriktionsschnittstellen im Vektor genau bekannt sind und da Eco RI bei den meisten oben beschriebenen Klonen den Vektor genau herausschneidet, lassen sich die Schnittstellen im klonierten Fragment von denen im Vektor sicher

unterscheiden. Durch den Vergleich der Fragmente zwischen Einfach- und Doppelverdau ist die Lokalisierung von bis zu zwei Schnittstellen im klonierten Fragment möglich. Ein Beispiel für diese Strategie ist in Abb. 12 für die Restriktionskartierung von Klon pDv-161 gezeigt. Dieses Gel wurde auch geblottet und nacheinander gegen AA/TT und GT/CA hybridisiert. Damit war es möglich, die einfachen Sequenzen bestimmten Regionen des Klons zuzuordnen.
Wenn mehr als zwei Schnittstellen eines Enzyms in dem klonierten Fragment gefunden wurden, wurde ein partieller Restriktionsverdau gemacht. Aus den dabei entstehenden Teilprodukten und mit der Vorinformation aus dem oben beschriebenen Doppelverdau läßt sich dann in der Regel die Reihenfolge der Fragmente zueinander finden.
Weitere Schnittstellen wurden noch durch andere geeignete Doppelverdaus, zum Teil auch mit isolierten Fragmenten gefunden.

Restriktionsenzyme, die nur vier Nukleotide als Erkennungssequenz besitzen, haben in größeren DNA-Fragmenten zu viele Schnittstellen (statistisch alle 256 bp), um mit den oben genannten Methoden einfach bestimmt werden zu können. Man kann in diesen Fällen eine, von Smith & Birnstiel (1976) entwickelte Partialverdaumethode anwenden. Dabei wird das betreffende DNA-Stück endmarkiert und mit einem Enzym, das nur einmal und möglichst nahe an einem Ende schneidet, nachverdaut. Das so markierte Fragment wird nun mit verschiedenen Enzymen partiell verdaut. Die radioaktiven Banden, die man nach der Autoradiographie des Gels dann findet, ergeben direkt die Reihenfolge und Länge der Restriktionsfragmente. Ein Beispiel dafür ist in Abb. 13 für die Restriktionskartierung des Klons pDv-1 gezeigt.

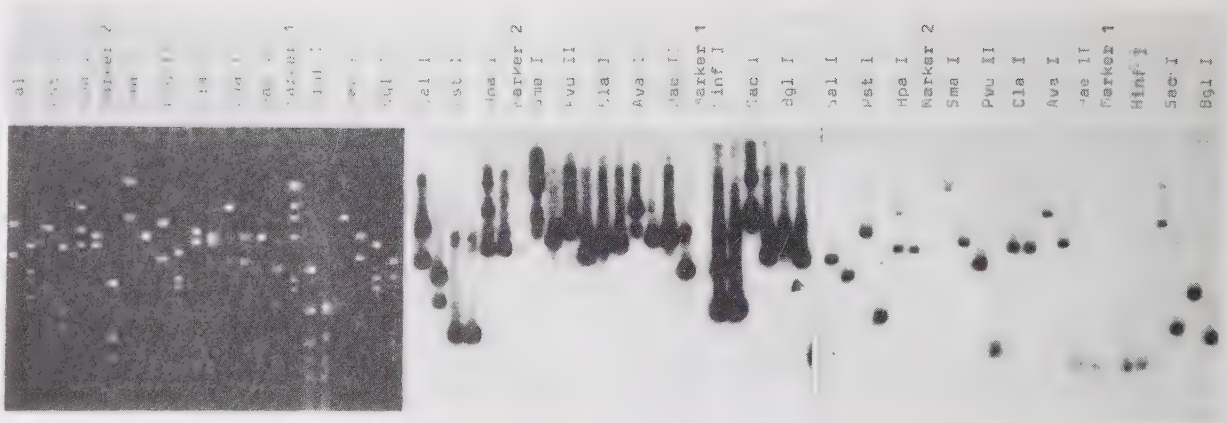


Abb. 12: Beispiel für die Restriktionskartierung und die Lokalisierung der simplen Sequenzen in Klon pDv-161.

Mit jedem Enzym wurde ein Einfach- und ein Doppelverdau durchgeführt, wobei die jeweils erste Spur den Einfachverdau, die jeweils zweite Spur den Doppelverdau mit diesem Enzym und Eco RI zeigt. Die geschnittenen DNAs wurden auf einem 0, 8% Agarosegel aufgetrennt. A = gefärbtes Gel; B = Hybridisierung mit AA/TT; C = Hybridisierung mit GT/CA. (Unterschiedliche Expositionszeiten).

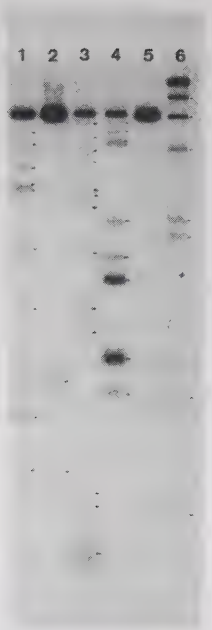


Abb. 13: Beispiel für die Restriktionskartierung nach der Partialverdaumethode, Klon pDv-1

Das Eco RI Insert aus Klon pDv-1 wurde endmarkiert und mit Ava I nachgeschnitten (vgl. Restriktionskarte).
Die Abbildung zeigt die Autoradiographie der auf einem 0, 8% Agarosegel aufgetrennten, partiell verdauten Fragmente. 1 = Verdau mit Hha I; 2 = Verdau mit Hae II; 3 = Verdau mit Hae III; 4 = Verdau mit Sau 3A; 5 = unverdautes Fragment; 6 = λ-Hind III Marker.

III.4 Restriktionskarten und Sequenzen

Im folgenden sind die Restriktionskarten der in III.2. aufgeführten Klone sowie die Sequenzen der daraus sequenzierten Fragmente angegeben. In diese Karten ist jeweils auch die Lokalisation der in diesen Klonen gefundenen simplen Sequenzen eingetragen (schwarze Blöcke in der obersten Reihe, die Numerierung stimmt mit der in den Sequenzbildern angegebenen überein). Die Pfeile am unteren Rand der Figuren zeigen welche Restriktionsschnittstellen für die Sequenzierung benutzt wurden. Die Länge der Pfeile zeigt wie weit die entsprechenden Gele noch sicher lesbar waren.

III.4.1. Klon pDv-1

pDv-1 ist ein in die Eco RI-Stelle von pBR 322 subkloniertes Fragment (vgl. III.2.). Die Hind III-Stelle des Vektors schließt sich an die rechte der gezeichneten Eco RI-Stellen an. Die Kartierung mit den Enzymen Hha I, Hae II, Hae III und Sau 3A wurde nach der in III.3. beschriebenen Partialverdaumethode vorgenommen. Aus diesem Grund reicht die Karte für diese Enzyme nur bis zu der AvaI-Stelle.

Die Basennumerierung in der Sequenz fängt etwa 15 nt nach der in der Restriktionskarte linken Eco RI-Stelle an. Die Numerierung wird fortgesetzt über die nicht sequenzierte Lücke, wobei der Abstand nur anhand der Länge der Restriktionsfragmente geschätzt wurde. Die Nukleotide in Position 703 und 2662/2663 konnten nicht gelesen werden und sind daher nur durch Punkte markiert. In diesem Klon wurden insgesamt sieben einfache Sequenzen, die nicht kürzer als 10 nt sind, gefunden, und zwar in der folgenden Reihenfolge:

1: Ein AA/TT-Stück mit 10 bp Länge, Position 37-46.

2: Ein GA/CT-Stück mit einer Länge von 57 bp (vier Substitutionen), Position 282-238.

3: Ein GTT/CAA-Stück mit einer Länge von 42 bp (zwei Substitutionen), Position 509-550.

4: Ein AT/TA-Stück mit einer Länge von 29 bp, Position 2340-2368.

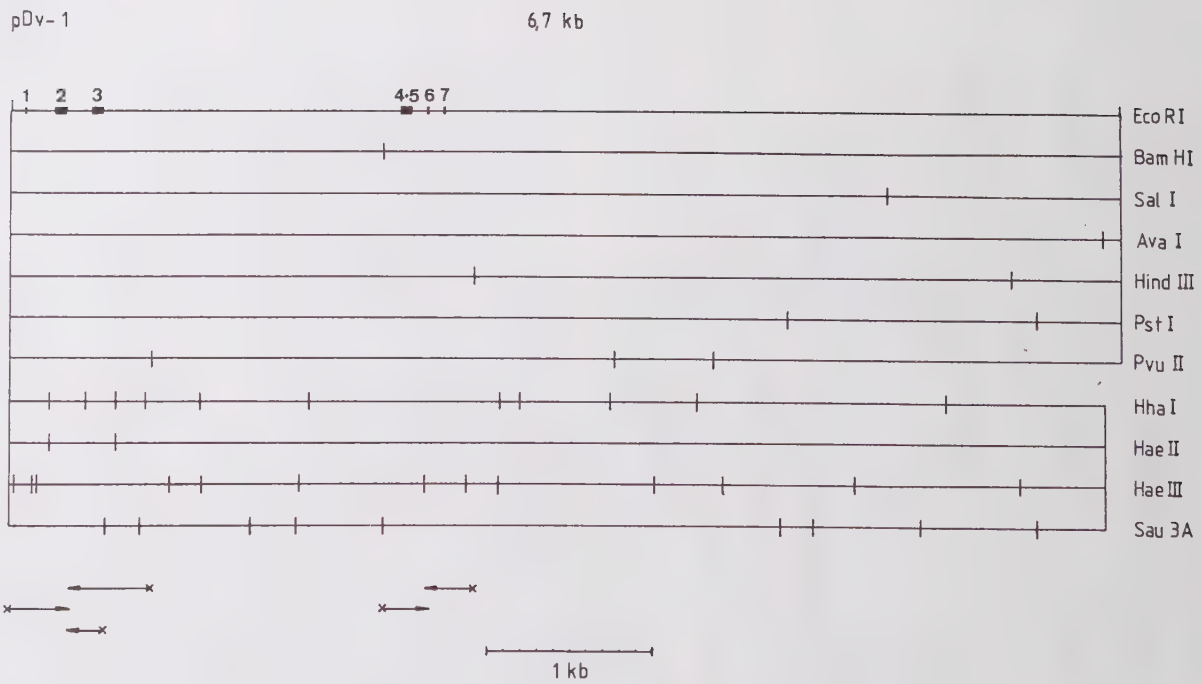
5: Ein GT/CA-Stück mit einer Länge von 25 bp (zwei Substitutionen), Position 2372-2396.

6: Ein GT/CA-Stück mit einer Länge von 15 bp (eine Substitution), Position 2486-2500.

7: Ein GT/CA-Stück mit einer Länge von 15 bp, Position 2529-2543.

Besonders interessant sind das AT/TA-Stück (4) und das GT/CA-Stück (5), die zeigen, daß unterschiedliche simple Sequenzen fast direkt ineinander übergehen können. Zusammen bilden diese Stücke eine Sequenz, die aus über 57 alternierenden Purin/Pyrimidin Nukleotiden besteht (vgl. Diskussion).

Es ist nicht ausgeschlossen, daß dieser Klon in den nicht sequenzierten Bereichen noch weitere simple Sequenzen enthält.


pDv-1
6,7 kb
1
2
3
4·5
6
7
EcoRI
Bam HI
Sal I
Ava I
Hind III
Pst I
Pvu II
Hha I
Hae II
Hae III
Sau 3A
1 kb

pDv-1

```
5' ...GGCCAACAAAAAATGTTTTTTTCTCATATAAATGCCAGCCGACGA    50
CGACTATTAAAATCCCATAATTTTTATTTACCAACCAAAAAAAAAAGTCC   100
                                          1
CAGCAAATGGGAAAAGGCTCATTAACATTAGCGGGTGGCAGCACCCACAG   150
ACACTCGGCCGCCATCCAGTGGCATTCAGGTTCAGTGTAGTCCACAATTT   200
CCCATGGAAGACCCCTCAATCCAAAGCAATCCAAGCGCCGCTCGGTGTGG   250
CTACTTAATGATAAAATATTTCATAACATACCCTCTCCCTCTCTCTCTCTC   300
TCTCTCTCTCTCTCTCTCTCTCCCTCTCTGTTTCTCCTTTGTTAAATG   350
          2
CCTTCCCGCTGACATGGACTTCTTAATTATTCAATGGACCCACTGAACGT   400
CGCATTGCTTCGATATCAAATACGACAACGGTACGAGTACAACATATGTT   450
TGCGCATTGCAAAAGTCTCTTCGTGTTGTAGAACTGCTGGCAGCAGGAGC   500
AGCAAAAGCAACAACAACAACAACAACAACAACAAAAACAACAATAACAA   550
                                  3
AAACTATAGTCTAGTGAGTCATAGAAGTTGATCGCCGAGTTCTACAAGCA   600
AACAGCGACTCGACATGTATCAAAGCGGCATTTGGCGCTTTGTTGTGGTT   650
CGAGGTAAATTGCCTTCGATTTTCACAAGTGCAAACGTGAACTATGGACA   700
AA.CTACAGGTGATGCAGGTATTTATAAATAGTTGATGTCGTAATGAGTA   750
CAGGGAAAAATTTCATTGAAAAGTGAGATCTCGATATGGGTAGAGAGCAC   800
ACATCATACGATTATTACGCGCATC...                ...AAT  2250
AAATCAATCGGTTTAGTTTTAAATAAAAGCCTGTGCCGGGCAACTCCGGG  2300
AGCCTCCTCCGTTACACACCAATACACAGTGCTTGCTCGTATATATATAT  2350
ATATATATATATATATATGTATGTGTGTGTGCGAGTGTGTGTGTGTAGCA  2400
4                          5
ATTCAGTAAAAAGTGCAAAGTGCAGCAAGACATTTGTGAACTTGGACATT  2450
TGTACGGCAACCTGCGATTTTCAAGTCGTAGTCGCCAGACACACACACAC  2500
                                          6
CAACACGAAAACGGCCGGACATTTGTTGACACACACACACACATGCATTC  2550
                                  7
TGTCTGAGTTAAATGGTCATTGGCATTTTTAGTGGCACTTGAAAGCCGTT  2600
AGCAAAGCAATTACAGCAAAAGACAACAACTTCAAGGCAGGTAAAAACGA  2650
CAGCAGACGAG..AATAAAAAGCACTATGCCACCTAATCCAGAGCTACCT  2700
GCTACCTGCCACATGTTGCAAGCATTATCCATAACGGGGCCATTGTGCAT  2750
GCAGC...
```

III.4.2. Klon pDv-142

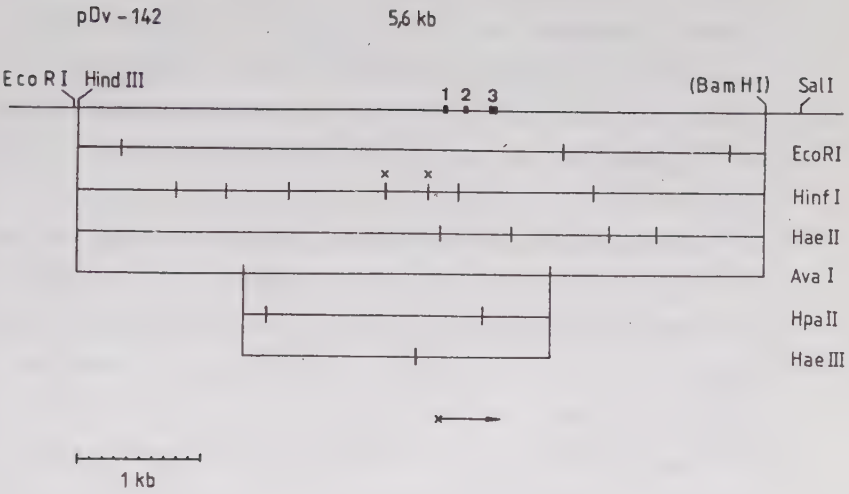
Bei pDv-142 ist ein Fragment zwischen die Hind III und die SalI-Stellen des Vektors pBR 322 subkloniert worden (vgl. III.2.). Das Stück Vektor zwischen der BamHI-Stelle und der SalI-Stelle wurde dabei regeneriert, das zwischen der HindIII-Stelle und der BamHI-Stelle wurde deletiert. Der Vektor ist damit um 0.35 kb verkürzt. Die rechts und links an das klonierte Fragment anschließenden Vektorschnittstellen sind in der Restriktionskarte mit eingetragen. Die BamHI-Stelle (eingeklammert) wird nicht mehr erkannt, da das klonierte Fragment nur mit einer Sau3A-Stelle endete (vgl. III.). Für HpaII und HaeIII wurden nur die Schnittstellen innerhalb des 2,5 kb AvaI-Fragments bestimmt. Die beiden mit einem 'x' bezeichneten HinfI-Stellen sind nicht sicher lokalisiert. Die drei davon betroffenen Fragmente können auch untereinander vertauscht sein. Die Basennumerierung in der Sequenz beginnt etwa 15 nt rechts von der in der Restriktionskarte linken HaeII-Stelle. In diesem Klon wurden drei GT/CA-Stücke und zwar in der folgenden Reihenfolge gefunden:

1: 24bp Länge (eine Substitution), Position 43-66.

2: 23bp Länge, Position 249-271.

3: 42bp Länge, (eine Substitution), Position 404-445.

Das erste GT/CA-Stück ist Teil einer 74nt langen Sequenz, die aus alternierenden Purin/Pyrimidin Basen besteht (zwei Substitutionen, gepunktete Linie). Der erste Teil dieser Sequenz hat eine simple Sequenzstruktur, bei der das Tetranukleotid CATA tandemartig wiederholt ist (Pfeile) mit zwei Substitutionen. Der zweite Teil dieser Sequenz hat keine offensichtliche simple Sequenzstruktur.



pDv-142

```
5' ..CGCAGACAAACCAGAGACATAGATACATGCATACATATGTGTGTG                    50
TGTGTTTGTGTGTGTGCACGTGTATATGTATGTGTATGCATGCAGCCCAA                   100
ACTGAAACCCACACACGCGCGAATTGTTGCGGCTCGATACCAAATCGAGT                   150
CTCAATTACCCGATAAACCCAAACACTCCACCCACAAGCCTCACAGCGGT                   200
CGGATCTCATTAGAGTGCGGGTTACGCAGCGGCAAAAGGGGATACGCATG                   250
TGTGTGTGTGTGTGTGTGTGTATTCGAATATTACTGCTGTGGTGTTTATA                   300
AAGCGTAACCGCCAGAGCCGGCAATTGGACCACGCCCGCTAGTTGTCCGA                   350
TCATTAAATGCTCAGCCAAATGAGCCATGCGAGAAATGCAGATGCCCACT                   400
CTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGCGTGTGTGTGTGATTGG                   450
GTGTGTGCCCG...
```

III.4.3. Klon pDv-161

pDv-161 ist ein, in die EcoRI-Stelle von pBR 322 kloniertes Fragment (vgl. III.2.). Die HindIII-Stelle des Vektors liegt neben der rechten gezeichneten EcoRI-Stelle. Auf die in der Restriktionskarte mit einem Fragezeichen bezeichnete HaeII-Stelle wurde nur indirekt geschlossen, sie ist daher nicht sicher.

Die Basennumerierung in der Sequenz beginnt 30 nt links von der in der Restriktionskarte linken PstI-Stelle. Sie setzt sich über die nicht sequenzierten Lücken fort, wobei die Abstände anhand der Länge der Restriktionsfragmente abgeschätzt wurden.

In diesem Klon wurden zehn Stücke mit einfachen Sequenzen gefunden. Stücke, die kürzer als zehn Nukleotide sind, wurden nicht berücksichtigt, obwohl auch davon mehrere zu finden sind, zum Beispiel um Position 124 oder Position 175 herum.

1: Ein GTT/CAA-Stück, 21 bp lang, zwei Substitutionen, Position 8-28.

2: Ein GT/CA-Stück, 23 bp lang, eine Substitution, Position 68-90.

3: Eine Struktur, ähnlich wie die, die in Klon pDv-191 (siehe III.4.4) gefunden wurde, nämlich alternierende C und T mit einer Unterbrechung durch 3 C; insgesamt 14 nt lang, Position 213-226.

4: Ein GTT/CAA-Stück, 20 bp lang, eine Substitution, Position 266-285.

5: Ein AA/TT-Stück, 23 bp lang, zwei Substitutionen, Position 366-388.

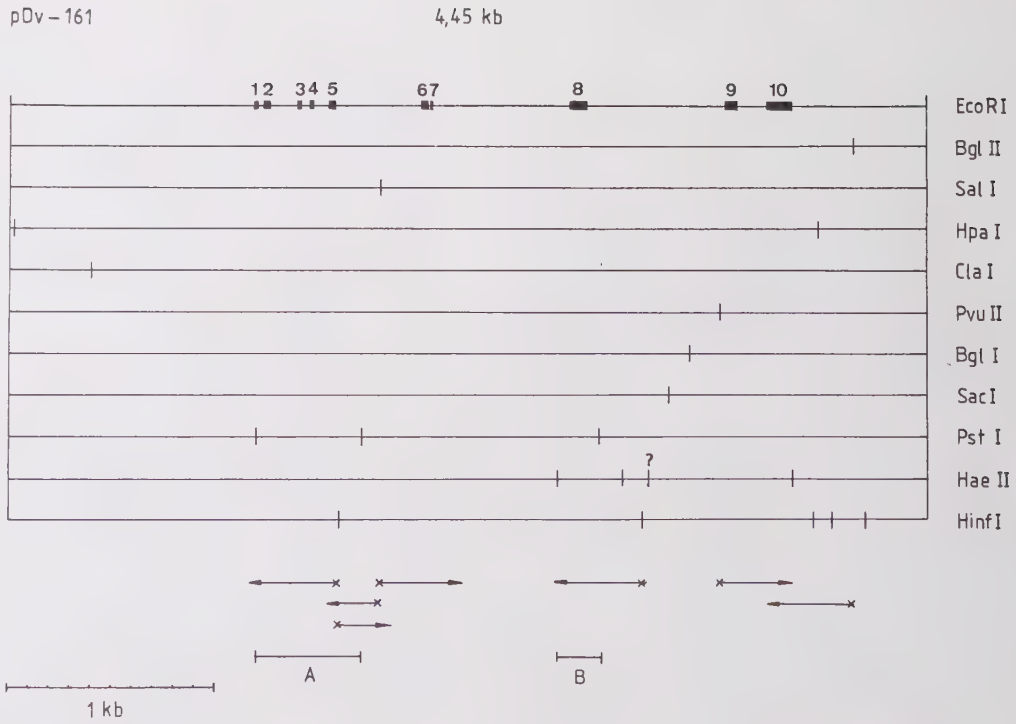
6: Ein AA/TT-Stück, 41 bp lang, fünf Substitutionen, Position 809-849.

7: Ein AA/TT-Stück, 10 bp lang, Position 854-863.

8: Ein CAG/GTC-Stück, 92 bp lang mit 12 Substitutionen, Position 1554-1645.

9: Ein Stück aus tandemartigen Wiederholungen des Hexanukleotides GCAACT, 55 nt lang mit zwei Substitutionen und zwei Insertionen (markiert mit 'x'), Position 2310-2364

10: Ein GT/CA Stück, 122 bp lang, zwei Substitutionen, Position 2507-2628.


pDv – 161
4,45 kb
1 2
3 4
5
6 7
8
9
10
EcoRI
Bgl II
Sal I
Hpa I
Cla I
Pvu II
Bgl I
Sac I
Pst I
?
Hae II
Hinf I
A
B
1 kb

pDv-161

```
5' ..CTCAACAGCAACAACAACGACAAACTGCAGAATGAAATTATTATG      50
                1
GTGATGGTGTTTAACCTACACACACACATACACACACACATATAGATACA     100
                          2
AATACAAACACAAACACGCACACAGACAACTAAAAACGTACACTGAAAAA     150
ATAAAATGAAAGTCTATCAACACACACACGGCTCTGTACCTTTTTACGTT     200
TGATTTCCTCTACTCCCCCTCTCTCTGTATATTGCCACACACTCCGCCAC     250
                    3
ACAACGGTAACGAGTACAACAACAATAACAACAACGAAAATTACATACAA     300
                          4
TACAATATTAATCCCAACAGTAACAAACAAGCGTTCTAAAAATAACTATA     350
ACAAAACAAAGAAGGAAAACAAAAAAAAAAAAACAAAATTTCAATTAAAA     400
                          5
AACAAAAGAGTATCATGAGTCAGTCAAAATTAAAGGAAAAAATTAAGCAC     450
ACATTTGTGATTGACTGTAACTATAACTGTAACGGTAACTGTAGCTCGAA     500
GCTGGCAGAAATCTTCAGCTAAAAACGTGATTAGCTTCGATCTGCAGCCA     550
ACCGCAGTATAGTTTAGATCGGATCACGACAGGATCAGGTCTATAGTCAG     600
CTATTGAAACGTCGACGACGTCATGCACGAACAATCAGAAGGGAAATAAA     650
TTTGTCTACTTGTTAATACAAAATTCTCTTATTTATTCACAATATATACT     700
ACTTATATAACAACGTGACTTGATCTCTCCTCCTTGAGAACCTCCTCAAC     750
TCTGACCCCCAATTTCATCTGATGCAAAAACAGTAGATAAATCAAAATCA     800
AAAACAGCAACAAAAAAAAAAAAAAATAAAAAACAAAAGAAATGAAAAAT     850
                          6
ACATTTTTTTTTTGTACGTAGGGCCACTTACCAAATTGTAATTTACCATA     900
    7
GATATCAAATAAATCCAAAGTTAAATTAAAAGTTAATA...  ...CAGC    1500
GCTGCACACGGACACGACACGCACACGACGACATACTTACACCCTCGCTG    1550
CCTCAACAACAGCAACAGCAGCAGCAACAACAGCAGCAGCAGCAGCAGCA    1600
                          8
GCAGCAGCACCCACAACAGCAGCAGCAGCACCAGCAACAACAGCACCAAT    1650
TGAGCAATGCAGGAAGCGGCAGTCAGTCAAGCAACGGTATTCTCGGCTGC    1700
AGTCCGCCCGCTGCGAACACACCAACAACGCGCAGCAGCAAATGTTCGCC    1750
GCGGTGATGGCCAATATGGCAGCCACGGCCAACGCTTCGGCCTCGACCAG    1800
CGGCAACAGCGCCAATAGCTCGCTGAGCAACTCGCAACTGAACACATCCC    1850
GGCAGCCTGAACGCCTCCAGCGGCGGCGATGATTTTCGC...
...CAAAAGGCAACTGCAACTGCGCAACAGCAACCGCAACTGCAACTGAA    2350
                  xx                  9          x
```

AACTGCAACTGCAATCACTCCTAGCTAAGGGCTATTCGCATAGTAACTAC 2400

AGATTAGATGAAACCTTTGTTAAAGTTTGTCTATTATATGAAAATGCGCG 2450

ATACACATTTGCAAATGACTGTAATCTTAATCAAATTAAAGCAACATTTA 2500

TACTTGCATACACACACACACACACACACACACACACACACACACACACA 2550

CACACACACACACACACACACACACACACACACACACACACACACACACA 2600

10

CACACACACACGCACACACACACACACAACGGCGCTTCTGCACGTATGTA 2650

TGTTTGTACAACAAAAGGCACATACATATCTGGGATATTCATGGTTTCCT 2700

GCAAACCTATACATATAGTAGATTTGGGTTGAGTCTCCAGAGAGCCCAGA 2750

GCCTGGTCTAGCTGTTAACTGTAGACATGCAATGTTAGTGTTTGAAAATG 2800

AAATGCATTTCATTAAGATTCATGCAACTCTGAATAACCAAAATTACACG 2850

TTTATATGCGAGACAAAATAATTACATTCTAGGCTAGTAACACTTAGTTG 2900

CCTTTTGTAGTTTAAGCAAATCAATAAAA...

III.4.4. Klon pDv-191

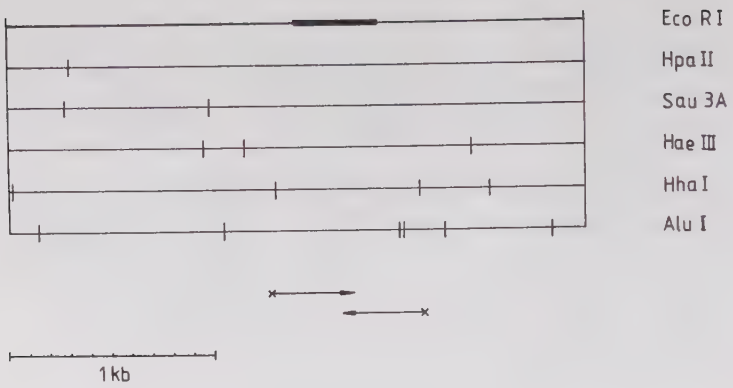
pDV-191 ist ein, in die EcoRI-Stelle von pBR 322 kloniertes Fragment (vgl. III.2.). Die HindIII-Stelle des Vektors schließt an die in der Restriktionskarte links gezeichnete Eco RI-Stelle an.

Zur Sequenzierung wurde das 0,7 kb HhaI-Fragment in die ClaI-Stelle von pBR322 subkloniert. Dazu wurden die überhängenden Enden von der ClaI-Stelle und den HhaI-Stellen mit Klenow-Polymerase glatt geschnitten (siehe Methoden), und die glatten Enden ligiert. Die HindIII-Stelle und die EcoRI-Stelle des Vektors wurden dann zum Sequenzieren verwendet. Die Orientierung der gezeigten Sequenz in dem ursprünglichen Klon ist daher nicht gesichert.

Die Basennumerierung in der Sequenz beginnt 5 nt links von der HhaI-Stelle. Das erste G der gezeigten Sequenz gehört also noch zu dieser Stelle.

Der Klon enthält eine simple Sequenzstruktur, die sich über ca. 400 bp erstreckt und die im wesentlichen aus alternierenden C und T besteht, die unregelmäßig von drei oder vier C unterbrochen wird. Am Anfang (Richtung 5´ Ende) wird sie viermal durch das Hexanukleotid GCCCCA unterbrochen. Unmittelbar vor ihr gibt es eine 36 nt lange, sich direkt wiederholende Sequenz (mit einer Substitution). Die ersten 5 nt dieser Sequenz erscheinen dann noch einmal bevor die simple Sequenz anfängt. Der wesentliche Bereich dieses Klons ist in dem Sequenzbild in Abb. 14 gezeigt. Die simple Sequenz und ihr Repetetionsmuster ist darin gut zu erkennen.

pDv 191 2,8 kb



pDv-191

```
5' ..GCAACGGAATGTCAACGATTTTGTCATTATGTAAGTGGCGACGTT      50
TCAGCCAAGTCTACCCCAACAAACCTGCTCTCCTGTACAGCCAAGTCTAC     100
CCCAACAAACCTGCTCTCATGTACAGCCTCTCTCTCTGCCCCATCTCTCT     150
GCCCCATCTCTCTCTGCCCCATCTCTCTCTGCCCCATCTCTCTCTCCCCC     200
CTCTCTCTCCCCCCTCTCTCTCTCCCCCCTCTCTCTCTCTCCCCCCTCTC     250
TCTCTCTCCCCCCTCTCTCTCTCTCTCTCCCCCTCTCTCTCTCCCCC      300
CCTCTCTCTCTCCCCCCTCTCTCTCCCCCCTCTCTCTCCCCCCTCTCTCT     350
CTCCCCCCTCTCTCTCTCCCCCTCTCTCTCCCCCTCTCTCTCTCCCCCTC     400
TCTCTCTCCCCCTCTCTCTCTCCCCCTCTCTCTCTCCCCCTCTCTCTCTC     450
CCCCCTCTCCATCTCCCCCTCTCTCTCCCCACGTCTCTCTCCCCCTTCTC     500
TCTCTCCCCCCTCTCTCTCGCTCCCTCTCTTTATTTTCCTCTCGCTCCCT     550
CTCTTCGAATGTTGATTTAAAAATGGGAAATTCAATAAAGTAAAAGCTCA     600
GAGCAACTTTCAAAGCTGAGCATCCTGTCATGCACTTTTCACCCTTTTAG     650
GGGTTAATTTCGCACACACATGAATGCATGCTCATCATGGTTGGCTGCCA     700
AAT...
```

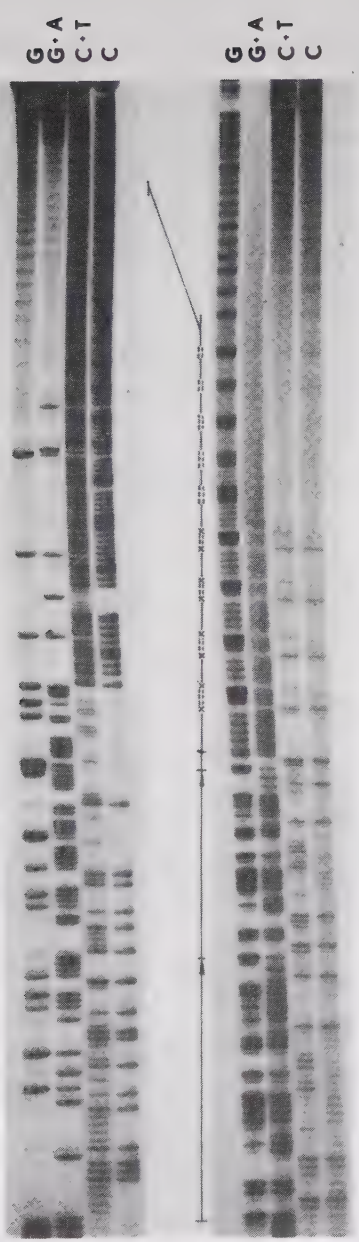


Abb. 14: Sequenzierungsgel aus Klon pDv-191.

Das 0,7 kb HhaI Fragment aus Klon pDv-191 wurde an beiden Seiten 3′-markiert und sesequenziert (vgl. Text).
Der linke Teil zeigt die Sequenz vom 3′-Ende her, der rechte Teil vom 5′-Ende her (jeweils bezogen auf die geschriebene Sequenz).
Die eingetragenen Markierungen und Symbole entsprechen denen im Sequenzbild.

III.4.5. Referenzklone pDv-182 und pMm-2

Die folgenden zwei Klone sind kurze, subklonierte DNA-Fragmente, die als Referenzklone für die in I.2. beschriebenen Versuche Verwendung fanden.

pDv-182

Ein 0,2 kb großes Stück aus Klon λ 3-161 wurde zwischen die EcoRI/HindIII Stellen von pUC 9 kloniert (vgl. III.2.). Die HindIII-Stelle wurde zum Sequenzieren verwendet. Die Basennumerierung beginnt also ca. 20 nt von dieser Stelle entfernt. Der Klon enthält nur eine simple Sequenz, nämlich ein 87 bp langes GT/CA-Stück (zwei Substitutionen).

pDv-182

5' ..AATAAAATGAAGCATACGAAAGCAGCGATCGCGCGCACACACTCA 50

CAGACACACACACACACACACACACACACACACACACACACACACACACACA 100

CACACACACACACACACACACACACACGTAATTCCAACTTTTTTTCCTGC 150

AACGCCGCGTGC...

pMm-2

Dieser Klon trägt ein Stück eines Plasmids, das das Histokompatibilitätsantigen H-2k^d der Maus enthält (Kvist et al., 1983) (erhalten von Sune Kvist). Das subklonierte Stück ist ein XbaI-Fragment aus dem dritten Intron dieses Gens, das nach Auffüllung der überhängenden Enden in die ClaI-Stelle von pBR 322 kloniert wurde. Die gezeigte Sequenz ist aus (Kvist et al., 1983) entnommen. Die Basennumerierung beginnt und endet an den XbaI-Schnittstellen.

Der Klon trägt ein GA/CT-Stück mit einer Länge von 44 bp und daran anschließend ein AA/TT-Stück mit 26 bp Länge (zwei Substitutionen).

pMm-2

5' ..ATTTTCCAAAGAATACATTCTCACAGATCCCTCCCTGTCTGTGGG 50

GTTTGCACCCCTTCAACAACCCAATTCTATCTATTCCTACAGTGGTGAAT 100

TGTCACATGAGGCCTTATGGGGTACCCTGGAGAAATATAAATCGTGGATT 150

TTCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 200

CGTTTTTTGTTTTTTGTTTTTTTTTTTTTTCTAG

III.5. Zusammenfassung der Sequenzierungsergebnisse

In den wenigen in dieser Arbeit sequenzierten Klonen wurden bereits eine ganze Reihe, zum Teil sehr verschiedener simpler Sequenzen gefunden. Sie zeigen eine starke Heterogenität, selbst innerhalb der gleichen Klasse. Besonders interessant sind die Fälle, wo zwei verschiedene simple Sequenzen praktisch direkt ineinander übergehen (z.B. in Klon pDv-1, aber auch in Klon pmM-2). Dies deutet bereits darauf hin, daß nicht irgendwelche bestimmte flankierende Sequenzen für die simplen Sequenzen notwendig sind, da man so etwas sonst bei allen beobachten müßte.

Beim Vergleich der flankierenden Regionen aller in dieser Arbeit gefundenen simplen Sequenzen (Abb. 15) konnten keine Homologien festgestellt werden. Dies schließt aber nicht aus, daß in der weiteren Umgebung dieser Elemente noch Homologien oder bestimmte Strukturen vorhanden sein könnten. Lediglich bei den beiden AA/TT-Elementen in Klon pDv-161 ist auffallend, daß auch die Umgebung dieser Elemente teilweise sehr A-reich ist (auf dem gezeigten Strang).

TATATATATATATATATATATATATATGTATGTG..(25)..GTGTAGCAATTCAGTAAAAAGTGCAAAGTGCAGC aus pDv-1 (5)
GTCAACAAATGTCCGGCCGTTTTCGTGTTGGTGT..(15)..TCTGGCGACTACGACTTGAAAATCGCAGGTTGCC aus pDv-1 (6)
CCAATGACCATTTAACTCAGACAGAATGCATGTG..(15)..GTGTCAACAAATGTCCGGCCGTTTTCGTGTTGGT aus pDv-1 (7)
AAACCAGAGACATAGATACATGCATACATATGTG..(24)..TGTGCACGTGTATATGTATGTGTATGCATGCAGC aus pDv-142 (1)
GGGTTACGCAGCGGCAAAAGGGGATACGCATGTG..(23)..GTGTATTCGAATATTACTGCTGTGGTGTTTATAA aus pDv-142 (2)
GCCATGCGAGAAATGCAGATGCCCACTCTATGTG..(42)..TGTGATTGGGTGTGTGCCCG aus pDv-142 (3)
TGCGTGTTTGTGTTTGTATTTGTATCTATATGTG..(23)..GTGTAGGTTAAACACCATCACCATAATAATTTCA aus pDv-161 (2)
TACAAACATACATACGTGCAGAAGCGCCGTTGTG..(123).TATGCAAGTATAAATGTTGCTTTAATTTGATTAA aus pDv-161 (10)
CGGCGTTGCAGGAAAAAAAGTTGGAATTACGTGT..(87)..TGTGCGCGCGATCGCTGCTTTCGTATGCTTCATT aus pDv-182

TCCATGTCAGCGGGAAGGCATTTAACAAAGGAGA..(57)..AGAGGGTATGTTATGAAATATTTATCATTAAGTA aus pDv-1 (2)
ATTCGAAGAGAGGGAGCGAGAGGAAAATAAAGAG..(403).AGAGGCTGTACATGAGAGCAGGTTTGTTGGGGTA aus pDv-191
AAAAAAAAAAAAAACAAAAAACAAAAAACGGAGA..(44)..GAGAAAAGAAAATCCACGATTTATATTTCTCCAG aus pMm-2

TTAAAATCCCATAATTTTTATTTACCAACCAAAA..(10)..AAAAGTCCCAGCAAATGGGAAAAGGCTCATTAAC aus pDv-1 (1)
CTAAAAATAACTATAACAAAACAAAGAAGGAAAA..(23)..AAAATTTCAATTAAAAAACAAAAGAGTATCATGA aus pDv-161 (5)
AACAGTAGATAAATCAAAATCAAAAACAGCAACA..(41)..AAAATACATTTTTTTTTTTGTACGTAGGGCCACTT aus pDv-161 (6)
AATTACAATTTGGTAAGTGGCCCTACGTACAAAA..(10)..AAAATGTATTTTTCATTTCTTTTGTTTTTTATTT aus pDv-161 (7)
GCCAGGGCTATACAGAGAAACCCTGTCTAGAAAA..(28)..AAAACGGAGAGAGAGAGAGAGAGAGAGAGAGA aus pMm-2

Abb. 15: Vergleich der flankierenden Sequenzen der in dieser Arbeit gefundenen GT/CA, GA/CT und AA/TT Elemente.

Jeweils 30 nt der Sequenz links und rechts der Elemente sind untereinander geschrieben, dazu jeweils die ersten vier nt und die letzten vier nt des Elements selbst. Die Zahlen in der Mitte geben die Länge des ganzen Elements an. Die Stränge sind so geschrieben, daß die vergleichbaren Elemente jeweils in der gleichen Richtung gelesen werden. Das 5´-Ende ist jeweils links.

III.6. Genomblothybridisierungen an D.virilis DNA

In den sequenzierten Klonen wurden zusätzlich zu den bereits durch Hybridisierung nachgewiesenen einfachen Sequenzen noch eine Reihe anderer simpler Sequenzen gefunden. Es war daher die Frage, ob auch diese repetetive Komponenten darstellen und ob etwa die gesamte klonierte DNA in diesen Bereichen repetetiv ist. Es wurde versucht, diese Fragen durch Hybridisierung geeigneter Fragmente an Genomblots von D.virilis zu klären.

III.6.1. Hybridisierung mit Klon pDv-192

Der Klon pDv-192 ist das zur Sequenzierung in pBR 322 subklonierte 0, 7 kb HhaI Fragment aus Klon pDv-191 (vgl. III.4.4.). Der Klon enthält im wesentlichen GA/CT als simple Sequenz, die jedoch oftmals unterbrochen ist. Es gibt darin kein zusammenhängendes GA/CT-Stück, das länger als 19 bp (bzw. 35 bp mit zwei Substitutionen) ist. Unter stringenten Hybridisierungs- und Waschbedingungen sollten daher die GA/CT-Anteile nicht mehr hybridisieren, da sie zu kurz sind, bzw. zu viele falsch gepaarte Bereiche enthalten würden (vgl. IV.1.). Außerdem wurde bei der Hybridisierung noch eukaryontische DNA (aus Heringsspermien) zugegeben, die mit simplen Sequenzen kompetiert (Miesfeld et al., 1981). Abbildung 16 (a + b) zeigt den Vergleich zwischen stringenter und nicht stringenter Hybridisierung des Klons an D.virilis EcoRI Genomblots. In dem ersten Fall ist nur eine Bande bei 2, 8 kb zu sehen, im zweiten ein Schmier und viele Banden, ähnlich wie bei der Hybridisierung von GA/CT an D.melanogaster Genomblots (vgl. Abb. 4). Dieses Ergebnis

zeigt, daß die Regionen, die in diesem Klon die simplen Sequenzen flankieren, nur einmal im Genom vorkommen, oder aber zumindest nicht oft repetiert sind. Mit diesem Experiment kann nämlich nicht ausgeschlossen werden, daß das hybridisierende EcoRI-Fragment selbst Teil eines längeren DNA-Stücks ist, das mehrmals (z.B. tandemartig) im Genom repetiert ist.

III.6.2. Hybridisierung mit Fragmenten aus Klon pDv-161

Aus diesem Klon wurden zwei Fragmente isoliert und nicktranslatiert. Das erste ist 0,5 kb lang und durch die PstI-Stellen an Position 34 und 542 begrenzt, das zweite ist 0,2 kb lang und wird durch die HaeII-Stelle an Position 1502 und durch die PstI-Stelle an Position 1697 begrenzt (vgl. Sequenz und die am unteren Rand gekennzeichneten Fragmente A und B in der Restriktionskarte). Das erste enthält vier verschiedene, aber auch sehr kurze simple Sequenzen, das zweite enthält die simple Trinukleotidsequenz CAG/GTC.
Wird das erste unter stringenten Bedingungen an einen D.virilis EcoRI-Genomblot hybridisiert (mit eukaryontischer Kompetetions DNA), erhält man nur eine Bande bei 4,4 kb (Abb. 16c). Dieses Ergebnis zeigt, daß auch hier die die simplen Sequenzen flankierenden Regionen nur einmal im Genom vorkommen (oder zumindest nicht stark repetiert sind, s.o.).
Das zweite Fragment wurde ebenfalls an einen D.virilis EcoRI - Genomblot hybridisiert (ohne eukaryontische Kompetetions DNA; mit stringentem Waschschritt). Es wurde ein Schmier über den gesamten Molekulargewichtsbereich erhalten, was zeigt, daß diese Sequenz ebenfalls hochrepetetiv ist und im Genom verteilt ist (Abb. 16d).

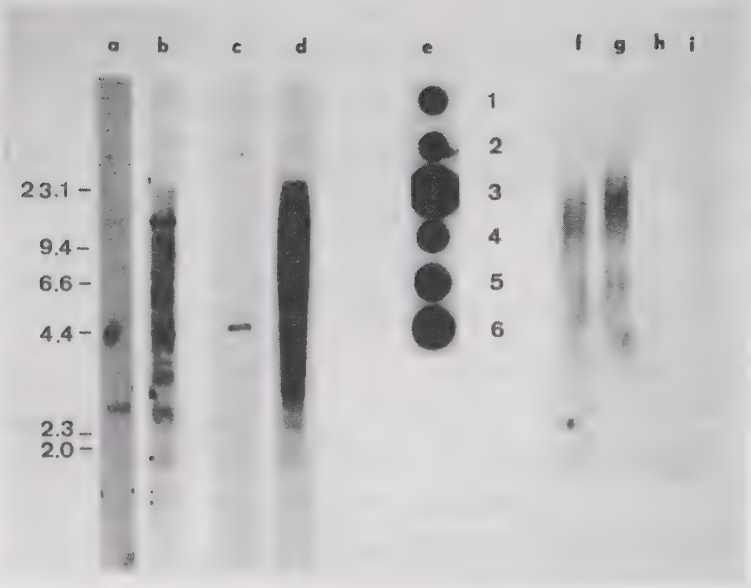


Abb. 16: Hybridisierungen mit Fragmenten aus den Klonen pDv-191 und pDv-161.

a-d = Genomblots von D.virilis, je 2µg EcoRI verdaute DNA pro Spur; e = Dotblots von DNA von sechs verschiedenen Eukaryonten, 1 = Mensch; 2 = Xenopus; 3 = Drosophila; 4 = Seeigel; 5 = Weizen; 6 = Hefe. f-i = RNA-Fraktionen aus Drosophila Kc-Zellen, f = Kern-RNA (ca. 5µg); g = cytoplasmatische RNA (ca. 15µg); h = poly-A^- RNA (ca. 5µg); i = poly-A^+ RNA (ca. 0,3µg).

a: hybridisiert unter stringenten Bedingungen mit Klon pDv-192 (siehe Text); b: wie a, aber nicht stringente Bedingungen; c: hybridisiert unter stringenten Bedingungen mit dem 0,5 kb PstI Fragment aus Klon pDv-161 (siehe Text); d-i: hybridisiert mit einem 0,2 kb PstI/HaeII Fragment (siehe Text), das die simple Sequenz CAG/GTC enthält.

Wird dieses Fragment an dot blots der anderen Eukaryonten hybridisiert, so erhält man bei allen starke Signale, (Abb. 16e), also einen Hinweis, daß sie auch bei diesen hoch repetetiv sind. Weiterhin kann man diese Sequenz auch in der RNA finden, wie Abbildung 16f-i zeigt. In diesem Versuch wurde das oben genannte Fragment gegen die fraktionierte Kc-Zell RNA (vgl. II.2.) hybridisiert. Ähnlich wie die simple Sequenz GT/CA hybridisiert es auch sehr stark mit der cytoplasmatischen und der poly-A^+ Fraktion, allerdings mit einem anderen 'Banden'-muster als GT/CA (vgl. Abb. 6).

IV. Diskussion der Experimente

IV.1. Hybridisierungsspezifität

Die ersten Hinweise für simple Sequenzen in Eukaryonten lieferten die dot blot Experimente. Bei dieser Art von Experiment stellt sich allerdings sofort die Frage nach der Spezifität der Hybridisierung, zumal auch die Hybridisierungsbedingungen sehr permissiv waren. Entstanden dabei also Artefakte? Wie lang muß eine Sequenz sein, um ein spezifisches Signal zu ergeben?
Welchen Anteil von falsch gepaarten Nukleotiden darf sie besitzen?

Hybridisierungsartefakte können weitgehend ausgeschlossen werden, da keine Hybridisierung der simplen Sequenzen an die DNA von E.coli oder an die DNA der Klonierungsvektoren pBR 322 und Phage λ gefunden wurde (vgl. Abb. 11), auch dann nicht, wenn keine E.coli DNA als carrier bei der Hybridisierung anwesend war. Die Lokalisierung der simplen Sequenzen auf den Restriktionsfragmenten war immer eindeutig (vgl. Abb. 12).

Die Hybridisierungsbedingungen waren so kalkuliert, daß Sequenzen kürzer als 15 bp nicht mehr hybridisieren sollten, da ihre Schmelztemperatur dann unter der Hybridisierungstemperatur liegt. Die Richtigkeit dieser Kalkulation konnte durch Experimente teilweise bestätigt werden. Zunächst waren alle durch Hybridisierung gefundenen simplen Sequenzen länger als 15 bp. Dagegen zeigten andere simple Sequenzstrukturen, die nur durch Sequenzierung gefunden worden waren, und die kürzer als 15 bp waren, keine Hybridisierung. Zum Beispiel

hybridisierte das Stück aus 7 aufeinanderfolgenden T in Klon pDv-182 (Position 139 - 145) nicht mit AA/TT. Oder das GA/CT Stück (3) in Klon pDv-161 mit einer Länge von 14 bp (zwei Substitutionen) zeigte keine Hybridisierung mit GA/CT. Dagegen ergab das GT/CA-Stück (2) in Klon pDv-161, das 23 bp lang ist, (eine Substitution) ein schwaches Hybridisierungssignal, das erst nach längerer Exposition sichtbar wurde (vgl. Fig. 17A).

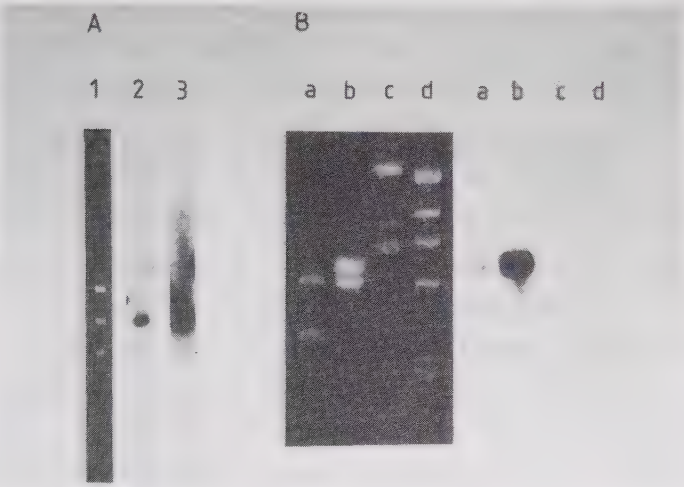


Abb. 17: Vergleich der Hybridisierungseffizienz von simplen Sequenzen unterschiedlicher Länge.

A: Klon pDv-161, geschnitten mit EcoRI und SalI und hybridisiert gegen GT/CA. 1 = gefärbtes Agarosegel; 2 = Hybridisierung mit GT/CA, kurze Exposition; 3 = wie 2, aber mit langer Exposition. Das Bild entspricht einem Ausschnitt aus Abb. 12.

B: Linke Seite gefärbtes Agarosegel, rechte Seite Hybridisierung mit GA/CT, mit stringentem Waschschritt. a = Klon pDv-191, Eco RI geschnitten; b = Klon 2.14 (der Klon aus dem pMm-2 subkloniert wurde), BamHI geschnitten; c = Klon λ X-54, EcoRI geschnitten; d = Marker, λ HindIII.

Andererseits können unter den gewählten Bedingungen sehr lange simple Sequenzen selbst dann noch hybridisieren, wenn sie einen sehr hohen Falschpaarungsanteil besitzen. So ist die Kreuzhybridisierung des Satelliten AAGAG aus D.melanogaster mit der simplen Sequenz GA/CT zu erklären (rechnerisch 50% Fehlpaarung). Eine rechnerische Fehlpaarung von 50% ergäbe sich auch bei einer Kreuzhybridisierung z.B. zwischen GA/CT und GT/CA, diese wurde jedoch nicht beobachtet.

Dennoch wurde für die weiteren Versuche sowie für die Quantifizierung der simplen Sequenzen in den Genomen (vgl. I.2.) noch ein zusätzlicher, stringenter Waschschritt eingeführt, der hybridisierende Teile, die kürzer als ca. 20 bp oder bei langen Stücken eine Fehlpaarung von mehr als 30% besitzen, wieder abschmelzen sollte. Auch diese rechnerisch ermittelten Bedingungen wurden durch ein Experiment kontrolliert. Der Klon pDv-191 zeigte nämlich nach dem stringenten Waschschritt eine mindestens 100-fach schwächere Hybridisierung mit GA/CT als Klon pMm-2 (Abb. 17B). pDv-191 trägt ein maximal 19bp langes ununterbrochenes GA/CT Stück, pMm-2 ein 43 bp langes.

Durch diesen Vergleich wird auch deutlich, daß die Quantifizierung der simplen Sequenzen in den Genomen stark von den angewandten Hybridisierungsbedingungen abhängt und daher nur einen groben Anhaltspunkt für die tatsächliche Situation liefert.

Für die Hybridisierung zwischen RNA und DNA gelten etwas andere Bedingungen. Insbesondere wenn Formamid in der Lösung ist, sind RNA-DNA Hybride stabiler als DNA-DNA Hybride (Casey & Davidson, 1977). Dieser Effekt ist allerdings von der

Basenzusammensetzung der Proben abhängig und nicht sehr genau kalkulierbar (Casey & Davidson, 1977). Daher wurden für die Versuche mit RNA die gleichen Bedingungen bei der Hybridisierung verwendet, wie für die DNA-Versuche, obwohl dabei in Betracht gezogen werden muß, daß die Spezifität etwas geringer ist. Dennoch sind diese Hybridisierungen nicht unspezifisch, wie z.B. Abbildung 11 zeigt. In diesem Versuch wurde radioaktiv markierte RNA an verschiedene DNA-Fragmente hybridisiert. Das Ergebnis war sehr spezifisch, wie sich auch nach der Sequenzierung einiger davon betroffener Klone ergab.

Auch die Northernblot Hybridisierungen sehen spezifisch aus, zumal sehr unterschiedliche Muster für die einzelnen simplen Polymere gefunden wurden (vgl. Abb. 6). Allerdings scheinen auch die ribosomalen Banden teilweise ein Signal zu ergeben. Dies ist möglicherweise unspezifisch und könnte mit der Menge der RNA an diesen Stellen zu tun haben (über 90% der aufgetragenen RNA-Menge konzentriert sich ja in diesen Banden).

Zusammenfassend kann man aber auf Grund aller oben genannten Argumente wohl davon ausgehen, daß die Hybridisierung mit simplen Sequenzen sowohl an DNA wie auch an RNA im wesentlichen spezifisch ist und kein Artefakt darstellt.

IV.2 Klonierungsbedingungen

Artefakte könnten auch durch die Klonierungsbedingungen entstehen. Das Ausgangsmaterial für die Mikroklonierung, nämlich die Quetschpräparate von Riesenchromosomen, wird ja kurzzeitig zur Fixierung einer Säurebehandlung unterzogen. Dies kann zu einer teilweisen Depurinisierung der DNA führen. Die betroffenen DNA-Stücke können dann entweder überhaupt nicht mehr kloniert werden oder müssen von dem E.coli eigenen Reparatursystem wieder regeneriert werden. Letzteres könnte zu Artefakten führen. Wenn bei diesem Prozeß allerdings regelmäßig simple Sequenzen entstehen sollten, müßte man in parallel klonierten Mikroklonbänken ähnliche Anteile davon finden, was nicht der Fall war (vgl. Tab. II).

Eine andere Frage ist die nach der Stabilität der simplen Sequenzen in den Klonierungsvektoren. Es gab keine Anzeichen, daß die Länge der simplen Sequenzen in den hier verwendeten Vektoren einer starken Änderung unterlag. Es wäre in diesem Fall nämlich nicht möglich gewesen, klare Sequenzbilder zu erhalten. Wenn sich die Länge der klonierten Sequenz nämlich in Subpopulationen des Klons ändert, würde man überlappende Sequenzbilder erhalten. Von anderen (Santoro & Costance, 1983) wurde dies für den einzelsträngigen Klonierungsvektor M 13 auch gefunden. Dagegen behielten alle drei in dieser Arbeit verwendeten Vektoren und Bakterienstämme die Sequenzen offenbar weitgehend stabil.

DISKUSSION

In der vorliegenden Arbeit konnte gezeigt werden, daß simple Sequenzen ein wesentlicher Bestandteil von eukaryontischen Genomen sind. Sie sind über das ganze Genom verteilt und werden auch transkribiert. Sie sind damit eindeutig von den einfachen Satellitensequenzen unterschieden.
Durch ihre simple Sequenz unterscheiden sie sich aber auch von anderen repetetiven Elementen. Man kann sie daher in eine neue Klasse einteilen, die Klasse der simplen repetetiven Sequenzen oder SRS-Elemente.

Aus den vorliegenden Sequenzdaten kann man einige Charakteristiken von SRS-Elementen ableiten.

1. SRS-Elemente sind kurz, meistens unter 100 bp lang, aber auch viel längere Stücke sind nicht ausgeschlossen (z.B. wie in Klon pDv-191). Bei sehr langen Stücken könnte man sie dann allerdings nicht mehr sicher von Satelliten-DNA abgrenzen. Auch eine Grenze für eine minimale Länge kann schwer angegeben werden, da sehr kurze Stücke von zufälligen Anordnungen nicht mehr zu unterscheiden sind. Ein 10 bp langes Stück einer simplen Sequenz sollte einmal unter 10^6 bp zufällig vorkommen, ein 15 bp langes einmal unter 10^9 (vgl. Rechnung in I.1.). Man kann also in diesem Bereich eine Grenze für die minimale Länge ziehen.
2. SRS-Elemente sind oft geclustert. In drei von vier näher analysierten Klonen wurden mehrere SRS-Elemente innerhalb von wenigen kb DNA gefunden. Zum Teil gehen sie unmittelbar ineinander über. Ähnliche Clusterungen wurden

auch von anderen gefunden (Höchtl & Zachau, 1983; Kvist et al., 1983).

3. Viele SRS-Elemente tragen Basensubstitutionen, die oft an Punktmutationen erinnern, also keine Phasenverschiebung in das Element hineinbringen.
4. SRS-Elemente besitzen keine Sequenzhomologien an ihren Enden. Bisher wurden keine gemeinsamen Sequenzstrukturen an den Grenzen von SRS-Elementen der gleichen Klasse oder unterschiedlicher Klassen gefunden, insbesondere keine kurzen, direkten repeats, die für andere repetetive Sequenzen oft typisch sind (Sharp, 1983). Andererseits ist nicht ausgeschlossen, aber auch nicht sehr wahrscheinlich, daß sehr lange Strukturen noch gefunden werden. Dagegen deuten die Versuche in III.6. darauf hin, daß die flankierenden Sequenzen unique sein können.
5. SRS-Elemente kommen in der Regel auch in der RNA vor.

Mit diesen Befunden stellen sich folgende Fragen:

1. Wie entstehen SRS-Elemente?
2. Warum kommen SRS-Elemente auch in der RNA vor?
3. Besitzen SRS-Elemente eine Funktion?

Zunächst sollen mögliche Entstehungsmechanismen diskutiert werden. Für die $SRS_{AA/TT}$ Elemente gibt es einen auf der Hand liegenden Mechanismus. Poly-A Stücke wurden nämlich in der Nähe von solchen Pseudogenen gefunden, die wahrscheinlich nach teilweiser Processierung ihrer mRNA wieder in das Genom integriert wurden, wobei auch der poly-A Schwanz der mRNA mitintegriert (Hollis et al., 1982; Sharp, 1983). Auch die repetetiven Sequenzen aus der Alu-Familie tragen an den Enden in der Regel oligo-A Stücke (Schmid & Jelinek, 1982). Auch bei ihnen gibt es Spekulationen, daß sie durch Reintegration

von (teilweise processierter) RNA in das Genom entstehen (Jagadeeswaran et al., 1981; Sharp, 1983). Ist so etwas auch als genereller Entstehungsmechanismus für SRS-Elemente vorstellbar? Eine Voraussetzung wäre dann, daß es in der Zelle Enzymaktivitäten gibt, die simple Sequenzen de novo polymerisieren können, so wie das für die poly-A Enden der mRNA der Fall ist. Über solche Aktivitäten ist wenig bekannt, aber es gibt offenbar Enzyme in der Zelle, die mit RNA und den Ribonukleotiden als Substrat Polymere bilden können (Mikoshiba et al., 1974; Hozumi et al., 1975; Milchev & Hadjiolov 1978). Über den Mechanismus und die dabei entstehenden Produkte ist allerdings nichts bekannt. Dagegen ist schon seit langem gezeigt, daß es in vitro mit Hilfe von DNA-Polymerase I oder RNA-Polymerase aus E.coli möglich ist, simple Polymere aus nur 4-10 nt langen Primern bzw. Templates zu polymerisieren (Wells et al., 1965; Byrd et al., 1965).

Mit einem Mechanismus, der von einer de novo Polymerisation simpler Sequenzen an RNA-Primer, cDNA-Transkription und Reintegration in das Genom ausgeht, wäre allerdings sehr schwer zu erklären, wie die teilweise enge Clusterung unterschiedlicher SRS-Elemente zu Stande kommen sollte.

Eine andere Entstehungsmöglichkeit für simple Sequenzen wäre eine sogenannte 'Slippage'-Reaktion. Sie wurde bereits 1966 in Zusammenhang mit der Entstehung von Frameshiftmutationen diskutiert (Streisinger et al., 1966). Während der DNA-Replikation oder während Reparaturvorgängen kann sich an Stellen, wo eine, oder mehrere Basen tandemartig wiederholt sind, das Basenpaarungsregister so verschieben, daß die entsprechende Stelle bei der Polymerisation um eine, oder

mehrere Repetetionseinheiten verlängert wird. Dadurch würden dann, insbesondere wenn dies mehrfach vorkommt, simple Sequenzen entstehen.
Viele, der in dieser Arbeit gefundenen Sequenzen könnten auf einen solchen Entstehungsmechanismus zurückgeführt werden. Die komplexe Struktur von Klon pDv-191 ist damit jedoch nicht zu erklären. Sie wäre jedoch durch einen anderen Mechanismus, nämlich unequal crossing over, erklärbar.

Unequal crossing over wurde von Smith (1976) vorgeschlagen als Mechanismus für die Entstehung von Satelliten DNA. Durch eine Computersimulation konnte er zeigen, daß aus einer Zufallssequenz, die einem ständigen unequal crosssing over (in dem Modell nur in kurzen Homologiebereichen) sowie zufälligen Punktmutationen unterliegt, auf die Dauer eine Sequenz entsteht, die nur noch aus sich tandemartig wiederholenden Einheiten besteht. In parallel durchgeführten Simulationen wurden dabei unterschiedliche, auch unterschiedlich lange repetierte Einheiten gefunden. Dies würde also bedeuten, daß auf die Dauer jede Sequenz, die nicht selektioniert wird, eine simple Sequenzstruktur annimmt, so lange sie unequal crossing over und Punktmutationen unterliegt. Dieses Konzept wurde zwar nur für Satelliten DNA entwickelt, ist aber auch sehr gut auf die Befunde in dieser Arbeit anwendbar. In leicht abgewandelter Form würde es dann lauten:

SRS-Elemente können in allen DNA-Regionen entstehen, die keiner oder keiner strengen Selektion unterliegen. Sie entstehen aus sehr kurzen, durch Zufall vorhandenen tandemartigen Wiederholungen von ein oder mehr Basen,

entweder durch einen 'Slippage'-Mechanismus oder durch unequal crossing over, wahrscheinlich aber durch ein Zusammenwirken von beiden.

Mit diesem Konzept ist leicht zu erklären, warum man zwar einerseits eine Verteilung der simplen Sequenzen über das ganze Genom findet, andererseits aber auch eine teilweise Clusterung. Nicht selektierte DNA-Regionen sind nämlich sehr wahrscheinlich über das ganze Genom verteilt, ebenso wie die Gene. Weiterhin könnten in größeren, nicht selektionierten Regionen theoretisch beliebig viele unterschiedliche SRS-Elemente, also eine Clusterung entstehen.

Wenn dieses Konzept zutrifft, sollte man finden, daß die Länge der SRS-Elemente einem starken Polymorphismus unterliegt. Drei Gründe sprächen dafür. Erstens entsteht bei unequal crossing over immer ein längerer und ein kürzerer Partner. Zweitens würden regelmäßige 'Slippage'-Reaktionen zu einer ständigen Verlängerung führen. Drittens ist bekannt, daß Sequenzen, die zwischen auf kurzer Distanz repetierten Regionen liegen, leicht deletiert werden können. Dafür ist wahrscheinlich ein dem 'Slippage'-Mechanismus ähnlicher Prozeß verantwortlich (Efstratiadis et al., 1980). Diesen drei Prozessen würden simple Sequenzen, unabhängig von ihrem Entstehungsmechanismus, auf jeden Fall unterliegen, es sei denn ihre Länge würde aus irgendeinem Grund selektioniert.
Ein solcher Längenpolymorphismus bei tandemartig wiederholten Sequenzen wurde auch bereits mehrfach beschrieben (Fedoroff 1979, Nikaido et al., 1981; Bell et al., 1982).

Wenn simple Sequenzen tatsächlich in jeder nicht selektionierten DNA-Region entstehen können, dann müssen sie auch in RNA vorkommen. Es liegt nämlich auf der Hand, daß es eine Reihe von transkribierten DNA-Regionen gibt, die keiner, oder keiner starken Selektion unterliegen. Zum Beispiel gibt es die Introns und die nicht translatierten Regionen in Genen. In einigen Introns würden ja auch SRS-Elemente gefunden (Slightom et al., 1980; Hamada et al., 1982a; Kvist et al., 1983).

In letzter Zeit konnte jedoch auch gezeigt werden, daß zumindest manche repetetiven Elemente eigene Promotoren besitzen, die wahrscheinlich auch in vivo zumindest teilweise erkannt werden und die Transkripte von dem betreffenden Element liefern (Haynes & Jelinek, 1981). Auf dieser Beobachtung basiert ja auch eine Theorie zur Ausbreitung von repetetiven Elementen im Genom (Jagadeeswaran et al., 1981). Fraglich ist, ob diese Transkripte nur zu dem letzteren Zweck entstehen und damit die Kriterien für selfish DNA (Doolittle & Sapienza, 1980; Orgel & Crick, 1980) erfüllen würden, oder ob sie eine Funktion für die Zelle darstellen. Die erste Version ist relativ wahrscheinlich, aber auch für die zweite Version wurde ein Beispiel gefunden, nämlich ein Transkript, das teilweise aus einem repetetiven Element der Alu-Familie besteht und das als 7S-RNA eine essentielle Funktion in der Zelle erfüllt (Ullu et al., 1982; Walter & Blobel, 1982). Geht man jedoch davon aus, daß die meisten dieser Transkripte nur 'selfish' sind, so wären diese 'Gene' gute Kandidaten dafür, ein SRS-Element zu erwerben und es zu cotranskribieren.

Im Falle der repetetiven Familie der MIF-Elemente konnte ein solcher Fall auf Sequenzebene gezeigt werden. Eines von zehn sequenzierten Elementen enthielt nämlich eine tandemartige Wiederholung eines Pentanukleotides, das wahrscheinlich durch unequal crossing over entstanden ist (Brown & Piechaczyk, 1983), ob dies allerdings auch mit transkribiert ist, wurde nicht untersucht. Aber von Promotoren solcher repetetiven Sequenzen können wahrscheinlich auch Transkripte starten, die über das Ende der repetetiven Sequenz hinausgehen und große Teile der benachbarten DNA-Sequenzen transkribieren, was für ein Alu-Element auch in vitro gezeigt wurde (Manley & Colozzo, 1982). Wenn dies auch in vivo passiert, könnte das natürlich auch zur Kotranskription von SRS-Elementen führen.

Keine der oben diskutierten Entstehungsmöglichkeiten würde SRS-Elementen eine Funktion zuordnen, weder in der RNA noch in der DNA. Eine indirekte Funktion wäre allerdings dadurch gegeben, daß sie wahrscheinlich als hot spots für Rekombinationsereignisse dienen. Möglicherweise entstehen sie ja gerade dadurch, daß im Bereich der simplen Sequenz die für die homologe Rekombination notwendige exakte Basenpaarung auch mit Rasterfehlern und damit leichter möglich ist. Es ist aber auch denkbar, daß mit Hilfe von SRS-Elementen heterologe DNA-Stücke miteinander homolog, nämlich im SRS-Element, rekombinieren können. Ein solcher Mechanismus trifft wahrscheinlich für das Rekombinationsereignis zu, das bei den Immunglobulingenen die Umschaltung von IgM auf IgG verursacht (Nikaido et al., 1981).
In einem speziellen Fall wurde auch bei menschlichen γ-Globingenen beobachtet, daß die simple Sequenz GT/CA Ausgangspunkt für ein Rekombinationsereignis, in diesem Fall

eine Genkonversion, war (Slightom et al., 1980). Das betreffende SRS-Element unterlag dabei übrigens einer starken Längenänderung, was die oben angesprochene Polymorphismushypothese bestätigt.

Einige der in dieser Arbeit diskutierten simplen Sequenzen können ungewöhnliche DNA-Strukturen bilden. GA/CT und GG/CC sind Purin-Pyrimidin Sequenzen und können möglicherweise viersträngige DNA-Helices bilden (Johnson & Morgan, 1978; Lee et al., 1980), auf jeden Fall aber zusammen mit einem Polypyrimidin-RNA Strang Tripelhelices (Morgan & Wells, 1968). Das gleiche gilt für AA/TT (Morgan & Wells, 1968). Dieser Umstand wurde dazu ausgenutzt, DNA-Fragmente, die AA/TT-Elemente enthalten, mit Hilfe einer poly-A Säule zu isolieren (Flavell et al., 1977; Sanders & Borst). Die DNA muß dazu nicht einmal denaturiert werden.
GT/CA und GC/CG sind Sequenzen mit alternierenden Purin/Pyrimidin Basen, die Z-DNA bilden können (Wang et al., 1979; Arnott et al., 1980), selbst unter physiologischen Bedingungen (Singleton et al., 1983; Nordheim & Rich, 1983). Da Z-DNA auch mit Hilfe von Antikörpern in biologischem Material (z.B. in Quetschpräparaten von Riesenchromosomen (Nordheim et al., 1981)) gefunden werden konnte, wurde spekuliert, ob die beiden oben genannten simplen Sequenzen dafür verantwortlich sind (Hamada et al., 1982b). Biologisch aktive Z-DNA wurde wahrscheinlich in den sogenannten Enhancersequenzen (Yaniv, 1982) von SV 40 entdeckt. Es gibt dort eine 8-10 nt lange Sequenz aus alternierenden Pu/Pyr Basen, die mit Antikörpern gegen Z-DNA reagiert. Wird diese Sequenz in vitro so mutiert, daß sie die Fähigkeit zur Z-DNA Bildung verliert, so verliert auch der Enhancer seine

Aktivität (Nordheim, pers. Kommun.). Der in diesem Zusammenhang wesentliche Punkt ist, daß diese Z-DNA keine simple Sequenz ist.

So gesehen ist es fraglich, ob die zur Z-DNA Bildung fähigen simplen Sequenzen auch biologische Funktionen erfüllen, zumal ihre Entstehung ja wahrscheinlich von Zufallsereignissen abhängig ist. Daher sind nicht-simple Sequenzen, die eine alternierende Pu/Pyr-Struktur aufweisen, in Zusammenhang mit biologischer Aktivität wahrscheinlich interessanter, da bei diesen eher möglich ist, daß sie auf Grund einer bestimmten Selektion zustande gekommen sind. Solch eine Sequenz wurde in einem in dieser Arbeit sequenzierten Klon auch tatsächlich gefunden (Klon pDv-141). In diesem Fall ist die simple Sequenz GT/CA in eine wesentlich längere Pu/Pyr-Struktur integriert. Man könnte dies so interpretieren, daß simple Sequenzen, obwohl sie zufällig entstehen, sekundär durchaus bestimmte Funktionen übernehmen könnten.

Das allgemeine Erscheinungsbild von simplen Sequenzen macht es jedoch wahrscheinlich, daß sie nicht dafür da sind, bestimmte Funktionen zu erfüllen, sondern als eine Art äußerste Form von selfish-DNA (Doolittle & Sapienza, 1980 ; Orgel & Crick, 1980) aufgefaßt werden können, die in ihrer Entstehung von Fehlern während der Replikation oder Rekombination abhängen.

ZUSAMMENFASSUNG

Durch Hybridisierung mit synthetischer DNA konnte gezeigt werden, daß alle möglichen DNA-Sequenzkombinationen, die aus einem Nukleotid oder aus zwei alternierenden Nukleotiden bestehen, als repetetive Komponenten in eukaryontischen Genomen vorkommen. Sie sind über das ganze Genom verteilt und werden auch in teilweise erheblicher Menge transkribiert. Zumindest in Drosophila kommen Transkripte dieser Sequenzen auch in poly-A^+ und cytoplasmatischer RNA vor.

Aus Drosophila virilis Genomklonbänken und aus zwei Mikroklonbänken wurden Klone isoliert, die simple Sequenzen enthalten. Diese wurden teilweise auf Sequenzebene charakterisiert. Aus diesen Versuchen ergab sich, daß simple Sequenzen sehr längenheterogen sind, aber meistens unter 100 bp lang sind. In ihren flankierenden Sequenzen konnten keine Homologien oder gemeinsame Strukturen festgestellt werden. Durch Sequenzierung wurden auch noch weitere simple Sequenzen gefunden, die aus drei, vier oder sechs tandemartig wiederholten Nukleotiden bestehen.

Es wird die Möglichkeit diskutiert, daß simple Sequenzen aus kurzen, im Genom zufällig vorhandenen Strukturen entweder durch einen 'Slippage'-Mechanismus oder durch unequal crossing over, wahrscheinlich durch beides, regelmäßig in nicht selektierten DNA-Regionen entstehen.

LITERATUR

Arnott, S., Chandrasekaran, R., Birdsall, D.L., Leslie, A.G.W. & Ratliff, R.L. (1980). Left-handed DNA helices. Nature (London) 283, 743-745.

Aviv, H. & Leder, P. (1972). Purification of biological active globin messenger RNA by chromatography on oligothymidilic acid cellulose. Proc. Nat. Acad. Sci. USA 69, 1408-1412.

Bell, G.I., Selby, M.J. & Rutter, W.J. (1982). The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. Nature (London) 295, 31-35.

Benton, W.D. & Davis, R.W. (1977). Screening λ gt recombinant clones by hybridization to single phages in situ. Science 196, 180-182.

Birnboim, H.C. & Straus, N.A. (1975). DNA from eucaryotic cells contains unusually long pyrimidine sequences. Can. J. Biochem. 53, 640-643.

Birnboim, H.C., Sederoff, R.R. & Paterson, M.C. (1979). Distribution of polypyrimidine-polypurine segments in DNA from divers organisms. Eur. J. Biochem. 98, 301-307.

Bishop, J.O., Rosbash, M. & Evans, D. (1974). Polynucleotide sequences in eucaryotic DNA and RNA that form ribonuclease-resistant complexes with polyuridylic acid. J. Mol. Biol. 85, 75-86.

Britten, R.J. & Kohne, D.E. (1968). Repeated Sequences in DNA. Science 161, 529-540.

Britten, R.J., Graham, D.E. & Neufeld, B.R. (1974). Analysis of repeating DNA sequences by reassociation. Methods in Enzymology, Vol. XXIX, 363-406; Academic Press, New York.

Brown, S.D.M. & Piechaczyk, M. (1983). Insertion sequences and tandem repetitions as sources of variation in a dispersed repeat family. J. Mol. Biol. 165, 249-256.

Brutlag, D.L. (1980). Molecular arrangement and evolution of heterochromatic DNA. Ann. Rev. Genet. 14, 121-144.

Byrd, C., Ohtsuka, E., Moon, M.W. & Khorana, H.G. (1965). Studies on polynucleotides XXXIX. Proc. Nat. Acad. Sci. USA 53, 79-86.

Calabretta, B., Robberson, D.L., Maizel, A.L. & Saunders, G.F. (1981). mRNA in human cells contains sequences complementary to the Alu family of repeated DNA. Proc. Natl. Acad. Sci. USA 78, 6003-6007.

Calos, M.P. & Miller, J.H. (1980). Transposable elements. Cell 20, 579-595.

Casey, J. & Davidson, N. (1977). Rates of formation and thermal stabilities of RNA: DNA duplexes at high concentrations of formamide. Nucl. Acids Res. 4, 1539-1552.

Chow, L.T. & Broker, T.R. (1981). Mapping RNA: DNA heteroduplexes by electron microscopy. In Electron microscopy in Biology, Vol. 1, 139-188; John Wiley & Sons, New York. Die einzelnen Formeln stammen aus folgenden Orginalarbeiten: GC-Abhängigkeit: Marmur & Doty, JMB 5, 109; Formamidabhängigkeit: Casey & Davidson, NAR, 4, 1539; Salzabhängigkeit: Dove & Davidson, JMB 5, 467; Längenabhängigkeit: Wetmur & Davidson, JMB 31, 349; Fehlpaarungsabhängigkeit: Smiley & Warner, NAR, 6, 1979.

Davidson, E.H. & Posakony, J.W. (1982). Repetitive sequence transcripts in development. Nature (London) 297, 633-635.

Denhardt, D.T. (1966). A membrane filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. 23, 641-646.

Deugau, K.V., Mitchel, R.E.J. & Birnboim, H.C. (1983). Nucleotide sequence of polypyrimidines form cloned mouse DNA as determined by basespecific blockage of exonucleaseaction. Analyt. Biochem. 129, 88-97.

Doolittle, W.F. & Sapienza, C. (1980). Selfish genes, the

phenotype paradigm and genome evolution. Nature (London) 284, 601-603.

Dubroff, L.M. & Nemer, M. (1975). Molecular classes of hetergeneous nuclear RNA in sea urchin embryos. J. Mol. Biol. 95, 455-476.

Dugaiczyk, A., Boyer, H.W. & Goodman, H.M. (1975). Ligation of Eco RI endonuclease-generated DNA fragments into linear and circular structures. J. Mol. Biol. 96, 171-184.

Duncan, R. & Humphreys, T. (1983). Oligo (U) sequences present in sea urchin maternal RNA decrease following fertilisation. Dev. Biol. 96, 258-262.

Echalier, A. & Ohanessian, B. (1970). In vitro 6, 162-172.

Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Conell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. & Proudfoot, N.J. (1980). The structure and evolution of the human γ-globin gene family. Cell 21, 653-668.

Epplen, J.T., McCarrey, J.R. Sutou, S. & Ohno, S. (1982). Base sequence of a cloned snake W-chromosome DNA fragment and identification of a male-specific putative mRNA in the mouse. Proc. Nat. Acad. Sci. USA 79, 3798-3802.

Fedoroff, N.V. (1979). On Spacers. Cell 16, 697-710.

Flavell, R.A., van den Berg, F.M. & Grosveld, G.C. (1977). Isolation and characterisation of the oligo(dA-dT) clusters and their flanking DNA segments in the rabbit genome. J. Mol. Biol. 115, 715-741.

Garoff, H. & Ansorge, W. (1981). Improvements of DNA sequencing gels. Analyt. Biochem. 115, 450-457.

Georgiev, G.P., Ilyin, Y.V., Chmeliauskaite, V.G., Ryskov, A.P., Kramerov, D.A., Skyabin, K.G., Krayev, A.S., Lukanidin, E.M. & Grigoryan, M.S. (1981). Mobile dispersed genetic elements and other middle repetitive DNA sequences in the genomes of Drosophila and mouse: Transcription and

biological significance. Cold Spring Harbour Symp. Quant. Biology, Vol. XLV-2, 641-653.

Glišin, W., Crkvenjakov, R. & Byus, C. (1974). Ribonucleic acid isolated by CsCl centrifugation. Biochemistry 13, 2633-2637.

Hamada, H., Petrino, M.G. & Kakunaga, T. (1982a). Molecular structure and evolutionary origin of human cardiac muscle actin gene. Proc. Nat. Acad. Sci. USA 79, 5901-5905.

Hamada, H., Petrino, M.G. & Kakunaga, T. (1982b). A novel repeated element with Z-DNA-forming potential is widely found in evolutionary diverse eucaryotic genomes. Proc. Nat. Acad. Sci. USA 79, 6465-6469.

Haynes, R.S. & Jelinek, W.R. (1981). Low molecular weight RNAs transcribed in vitro by RNA polymerase III from Alu-type dispersed repeats in chineese hamster DNA are also found in vivo. Proc. Nat. Acad. Sci. USA 78, 6130-6134.

Höchtl, J. & Zachau, H.G. (1983). A novel type of a aberrant recombination in immunoglobulin genes and its implication for V-J joining mechanism. Nature (London) 302, 260-263.

Hollis, G.F., Hieter, P.A., McBride, O.W., Swan, D. & Leder, P. (1982). Processed genes: a dispersed human immunoglobulin gene bearing evidence of RNA-type processing. Nature (London) 296, 321-325.

Hozumi, N., Haruna, I., Watanabe, I., Mikoshiba, K. & Tsukada, Y. (1975). Poly(U) polymerase in rat brain. Nature (London) 256, 337-339.

Jacobson, A., Firtel, R.A. & Lodish, H.F. (1974). Transcription of polydeoxythymidylate sequences in the genome of the cellular slime mold, Dictyostelium discoideum. Proc. Nat. Acad. Sci. USA 71, 1607-1611.

Jagadeeswaran, P., Forget, B.G. & Weissman, S.M. (1981). Short interspersed repetitive DNA elements in eucaryotes: Transposable DNA elements generated by reverse transcription of RNA Pol III transcripts? Cell 26, 141-142.

Johnson, D. & Morgan, R. (1978). Unique structures formed by pyrimidine-purine DNAs which may be four-stranded. Proc. Nat. Acad. Sci. USA 75, 1637-1641.

Kafatos, F.C., Jones, C.W. & Efstratiadis, A. (1979). Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridisation procedure. Nucl. Acids Res. 7, 1541-1552.

Karn, J., Brenner, S., Barnett, L. & Cesareni, G. (1980). Novel bacteriophage cloning vector. Proc. Nat. Acad. Sci. USA 77, 5172-5176.

Katinakis, P.K. & Burdon, R.H. (1981). Isolation and characterisation of a non-polyadenylated mRNA species with an affinity for poly (adenylic acid) from friend leukemia cells. Biochim. Biophys. Acta 653, 27-38.

Klenow, M. & Henningsen, I. (1970). Selective elimination of the exonuclease activity of the deoxribonucleic acid polymerase from E.coli by limited proteolysis. Proc. Nat. Acad. Sci. USA 65, 168-175.

Kominami, R., Muramatsu, M. & Moriwaki, K. (1983). A mouse type 2 Alu sequence (M2) is mobile in the genome. Nature (London) 301, 87-89.

Kvist, S., Roberts, L. & Dobberstein, B. (1983). Mouse histocompatibility genes: structure and organisation of a K^d gene. EMBO Journal 2, 245-254.

Langer, P.R., Waldrop, A.A. & Ward, C.D. (1981). Enzymatic synthesis of biotin-labelled polynucleotides: Novel nucleic acid affinity probes. Proc. Nat. Acad. Sci. USA 78, 6633-6637.

Lee, J.S., Evans, D.H. & Morgan, A.R. (1980). Polypurine DNAs and RNAs form secondary structures which may be tetra-stranded. Nucl. Acids Res. 8, 4305-4320.

Lehrach, H., Diamond, D., Wozney, J.M. & Boedtker, H. (1977). RNA molecular weight determinations by gelelectrophoresis

under denaturing conditions, a critical reexamination. Biochemistry 16, 4743-4751.

Lewin, B. (1980). Gene Expression 2 (Eucaryotic Chromosomes). John Wiley & Sons, New York.

Loening, U.E. (1967). The fractionation of high-molecular-weight ribonucleic acid by polyacrylamide-gel electrophoresis. Biochem. J. 102, 251-257.

Maizels, N. (1976). Dictyolstelium 17S, 25S, and 5S rDNAs lie within a 38 kb repeated unit. Cell 9, 431-438.

Maniatis, T., Fritsch, E.F. & Sambrook, I. (1982). Molecular cloning, a laboratory manual. Zu beziehen durch: Cold Spring Harbour Laboratory, Box 100, Cold Spring Harbour, New York.

Manley, J.L. & Colozzo, M.T. (1982). Synthesis in vitro of an exceptionally long RNA transcript promoted by an AluI sequence. Nature (London) 300, 376-379.

Maxam, A.M. & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods in Enzymology, Vol. LXV, 499-560; Academic Press, New York.

Miesfeld, R., Krystal, M. & Arnheim, N. (1981). A member of a new repeated sequence family which is conserved throughout eucaryotic evolution is found between the human α and β globin genes. Nucl. Acids. Res. 9, 5931-5947.

Mikoshiba, K., Tsukada, Y., Haruna, I. & Watanabe, I. (1974). RNA dependent RNA synthesis in brain. Nature (London) 249, 445-447.

Milchev, G.I. & Hadjiolov, A.A. (1978). Association of poly (A) and poly(U) Polymerases with cytoplasmic ribosomes. Eur. J. Biochem. 84, 113-121.

Mol, J.N.M., Flavell, R.A. & Borst, P. (1976). The presence of $(dA\text{-}dT)_{20\text{-}25}$ tracts in the DNA of primitive eucaryotes. Nucl. Acids Res. 3, 2367-2377.

Molloy, G.R., Thomas, W.L. & Darnell, J.E. (1972). Occurance of uridylate-rich oligonucleotide regions in heterogeneous

nuclear RNA of HeLa cells. Proc. Nat. Acad. Sci. USA 69, 3684-3688.

Morgan, A.R. & Wells, R.D. (1968). Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. J. Mol. Biol. 37, 63-80.

Murray, N.E., Brammar, W.J. & Murray, K. (1977). Lamboid phages that simplify the recovery of in vitro recombinants. Molec. Gen. Genet. 150, 53-61.

Murray, N.E. (1983). Phage λ and molecular cloning. Review in: The Bacteriophage Lambda Vol II; in Druck.

Nikaido, T., Nakai, S. & Honjo, T. (1981). Switch region of immunoglobulin Cμ gene is composed of simple tandem repetitive sequences. Nature (London) 292, 845-848.

Nishioka, Y. & Leder, P. (1980). Organisation and complete sequence of identical embryonic and plasmacytoma kV-region genes. J. Biol. Chem. 255, 3691-3694.

Nordheim, A., Pardue, M.L., Lafer, E.M., Möller, A., Stollar, B.D., & Rich, A. (1981). Antibodies to left-handed Z-DNA bind to interband regions of Drosophila polytene chromosomes. Nature (London) 294, 417-422.

Nordheim, A. & Rich, A. (1983). The sequence $(dC\text{-}dA)_n$- $(dG\text{-}dT)_n$ forms left-handed Z-DNA in negatively supercoiled plasmids. Proc. Nat. Acad. Sci. USA 80, 1821-1825.

Orgel, L.E. & Crick, F.H.C. (1980). Selfish DNA: the ultimate parasite. Nature (London) 284, 604-607.

Pardue, M.L. & Gall, J.G. (1975). Nucleic acid hybridization to the DNA of cytological preparations. In Methods in cell biology, Vol. 10, 1-16; Academic Press, New York.

Riede, I. (1982). DNA-Homologie und die somatische Paarung von Riesenchromosomen in Hybriden der D.virilis Gruppe. Dissertation, Universität Tübingen.

Rigby, P.W., Dieckmann, M., Rhodes, C. & Berg, P. (1977). Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol 113, 237-251.

Sanders, J.P.M. & Borst, P. (1977). The organization of genes in yeast mitochondrial DNA. Molec. gen. Genet. 157, 263-269.

Santoro, C. & Costanzo, F. (1983). Stretches of alternating poly (T-dG), with the capacity to form Z-DNA, are present in human liver transcripts. FEBS Letters, 155, 69-72.

Scalenghe, F., Turco, E., Edström, J.E., Pirotta, V. & Melli, M. (1981). Microdissection and cloning of DNA form a specific region of Drosophila melanogaster polytene chromosomes. Chromosoma 82, 205-216.

Schaffner, W., Kunz, G., Daetwyler, H., Telford, J., Smith, H.O. & Birnstiel, M.L. (1978). Genes and spacers of cloned sea urchin histone DNA analysed by sequencing. Cell 14, 655-671.

Schmid, C.W., Jelinek, W.R. (1982). The Alu family of dispersed repetitive sequences. Science 216, 1065-1070.

Sederoff, R., Lowenstein, L. & Birnboim, H.C. (1975). Polypyrimidine segments in Drosophila melanogaster DNA: II. Chromosome location and nucleotide sequence. Cell 5, 183-194.

Sharp, P.A. (1983). Conversion of RNA to DNA in mammals: Alu-like elements and pseudogenes. Nature (London) 301, 471-472.

Shenkin, A. & Burdon, R.H. (1972). Deoxyadenylate-rich sequences in mammalian DNA. FEBS Letters 22, 157-160.

Shenkin, A. & Burdon, R.H. (1974). Deoxyadenylate-rich and deoxyguanylate-rich regions in mammalian DNA. J. Mol. Biol. 85, 19-39.

Singer, M.F. (1982). SINEs and LINEs: Highly repeated short

and long interspersed Sequences in Mammalian Genomes. Cell 28, 433-434.

Singleton, C.K., Klysik, J. & Wells, R.D. (1983). Conformational flexibility of junctions between contigous B- and Z-DNAs in supercoiled plasmids. Proc. Nat. Acad. Sci. USA 80, 2447-2451.

Slightom, J.L., Blechl, A.E. & Smithies, O. (1980). Human fetal $^{G}\gamma$- and $^{A}\gamma$- globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell 21, 627-638.

Smith, G. (1976). Evolution of repeated DNA sequences by unequal crossing over. Science 191, 528-535.

Smith, H.O. & Birnstiel, M.L. (1976). A simple method for DNA restriction site mapping. Nucl. Acids Res. 3, 2387-2398.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gelelectrophoresis. J. Mol. Biol. 98, 503-517.

Straus, N.A. & Birnboim, H.C. (1974). Long pyrimidine tracts of L-cell DNA: Localisation to repeated DNA. Proc. Nat. Acad. Sci. USA 71, 2992-2995.

Straus, N.A. & Birnboim, H.C. (1976). Polypyrimidine sequences found in eucaryotic DNA have been conserved during evolution. Biochim. Biophys. Acta 454, 419-428.

Streisinger, G., Okada, Y. Emrich, J., Newton, J., Tsugita, A., Terzaghi, E. & Inouye, M. (1966). Frameshift mutations and the genetic code. Cold Spring Harbour Symp. Quant. Biol. 31, 77-84.

Sures, I., Lowry, J. & Kedes, L.H. (1978). The DNA sequence of sea urchin (S. purpuratus) H2A, H2B and H3 Histone coding and spacer regions. Cell 15, 1033-1044.

Sutcliffe, J.G. (1978). pBR 322 restriction map marked from the DNA sequence: Accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. 5, 2721-2728.

Tautz, D. & Renz, M. (1983). An optimized freeze squeeze method for the recovery of DNA-fragments from agarose gels. Analyt. Biochem. in press.

Ullu, E., Murphy, S. & Melli, M. (1982). Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an Alu sequence. Cell 29, 195-202.

Walter, P. & Blobel, G. (1982). Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmatic reticulum. Nature (London) 299, 691-698.

Wang, A. H-J., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Marel, G. & Rich A. (1979). Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature (London) 282, 680-686.

Wells, R.D., Ohtsuka, E. & Khorana, H.G. (1965). Studies on polynucleotides L. J. Mol. Biol. 14, 221-240.

Yaniv, M. (1982). Enhancing elements for activation of eucaryotic promoters. Nature (London) 297, 17-18.

Meine akademischen Lehrer waren

Botanik	Butterfaß, Kranz, Schaub
Zoologie	Ammermann, Engels, Halbach, Neuweiler, Wiltschko
Mikrobiologie	Braun, Bessler, Handtke, Zähner
Genetik	Forkmann, Hemleben, Lange, Meyer, Seyffert, Sperlich
Biochemie	Bisswanger, Entian, Fasold, Probst
Chemie	König, Rosemund
Physik	Barth, Brehm, Martiensen

LEBENSLAUF

Diethard Tautz

geboren am	17.8.57 in Glonn/Oberbayern
Eltern	Wilhelm Tautz (Aufzugmonteur) Eva-Maria Tautz, geb. Nentwig
Schulbesuch	1963 - 1967 Grundschule Grafing 1967 - 1976 Gymnasium Grafing Mai 1976 Abitur
Studium	1976 - 1978 Grundstudium der Biologie an der Universität Frankfurt/Main September 1978 Diplomvorprüfung 1978 - 1981 Hauptstudium der Biologie an der Universität Tübingen Mai 1981 Diplom im Fach Biologie, Hauptfach Genetik
Diplomarbeit	1981 am Max-Planck Institut für Biologie in Tübingen (Abteilung Beermann) unter Anleitung von Dr. habil. Manfred Renz. "Untersuchungen an Speicheldrüsen-Sekretproteinen von Drosophila virilis Arten".
Promotion	Von Juni 1981 bis März 1982 am Max-Planck Institut für Biologie in Tübingen (Abteilung Beermann) und von April 1982 bis Juni 1983 am Europäischen Laboratorium für Molekularbiologie in Heidelberg unter Anleitung von Dr. habil. Manfred Renz.

Druck: Polyfoto - Dr. Vogt KG
Forststraße 58, 7000 Stuttgart 1
Telefon (0711) 61 06 41

Druck: Polyfoto - Dr. Vogt KG
Forststraße 58, 7000 Stuttgart 1
Telefon (0711) 61 06 41